MICROCOPY RESOLUTION TEST CHART

NATIONAL BUREAU OF STANDARDS 1963 A

USCG-D-08-82

CAORF 50-7915-0

⑭

SIMULATORS FOR MARINER TRAINING AND LICENSING

PHASE 2: INVESTIGATION OF SIMULATOR CHARACTERISTICS FOR TRAINING SENIOR MARINERS

OCTOBER 1981

DTIC
ELECT
MAY 24 1982
S
E

Prepared for

U.S. DEPARTMENT OF COMMERCE
UNITED STATES MARITIME ADMINISTRATION
Office of Research and Development
Washington, D.C. 20230

and

8 2 0 5 2 4 0 5 0

FILE COPY

DTIC

U.S. DEPARTMENT OF TRANSPORTATION
UNITED STATES COAST GUARD
Office of Research and Development
Washington, D.C. 20590

LEGAL NOTICE

LEGAL NOTICE

BIBLIOGRAPHIC DATA SHEET	1. Report No. USCG-D-08-82	2.	3. Recipient's Accession No.

4. Title and Subtitle	5. Report Date
Simulators for Mariner Training and Licensing Phase 2: Investigation of Simulator Characteristics for Training Senior Mariners	October 1981
	6.

7. Author(s)	8. Performing Organization Report No.
T. J. Hammell, J. W. Gynther, J. A. Grasso, and M. E. Gaffney	CAORF 50-7915-02

9. Performing Organization Name and Address	10. Project/Task/Work Unit No.
National Maritime Research Center, Kings Point, New York 11024 *Computer Aided Operations Research Facility*	11. Contract/Grant No.

12. Sponsoring Organization Name and Address	13. Type of Report & Period Covered
Office of Research and Development Maritime Administration U.S.Dept. of Commerce Washington, D.C. 20230 — Office of Research and Development U.S.Coast Guard U.S.Dept. of Transportation Washington, D.C. 20590	Technical Report Research Study
	14.

16. Abstracts

As a result of the Phase 1 investigation, many specific gaps were identified in the empirical research literature regarding the use of simulators for mariner training. These gaps cover a wide range of variables relative to simulation and its influence upon training effectiveness and performance validity. Phase 2 clarifies several of the more important issues facing the design and use of simulators for mariner training by initiating a systematic investigation of several simulation and training program variables. The variables investigated during this phase of the research were: color/black and white visual scene, day/night simulation, horizontal field of view, target controllability, feedback methodology, and instructor differences.

17. Key Words and Document Analysis.   17a. Descriptors

Simulator
Simulator design issues
Simulator variables
Masters simulator-based training
Training objectives
Training effectiveness
Performance measures

Training variables
Training assistance technology
Training methodology
Screening process
Fractional factorial designs

17b Identifiers/Open-Ended Items

17c. COSATI Field/Group

18. Availability Statement	19. Security Classification (This Report) UNCLASSIFIED	21. No. of Pages 166
Approved for Release NTIS Springfield, Virginia	20. Security Classification (This Page) UNCLASSIFIED	22. Price

i

METRIC CONVERSION FACTORS

Approximate Conversions to Metric Measures

Symbol	When You Know	Multiply by	To Find	Symbol
LENGTH				
in	inches	*2.5	centimeters	cm
ft	feet	30	centimeters	cm
yd	yards	0.9	meters	m
mi	miles	1.6	kilometers	km
AREA				
in²	square inches	6.5	square centimeters	cm²
ft²	square feet	0.09	square meters	m²
yd²	square yards	0.8	square meters	m²
mi²	square miles	2.6	square kilometers	km²
	acres	0.4	hectares	ha
MASS (weight)				
oz	ounces	28	grams	g
lb	pounds	0.45	kilograms	kg
	short tons (2000 lb)	0.9	tonnes	t
VOLUME				
tsp	teaspoons	5	milliliters	ml
tbsp	tablespoons	15	milliliters	ml
fl oz	fluid ounces	30	milliliters	ml
c	cups	0.24	liters	l
pt	pints	0.47	liters	l
qt	quarts	0.95	liters	l
gal	gallons	3.8	liters	l
ft³	cubic feet	0.03	cubic meters	m³
yd³	cubic yards	0.76	cubic meters	m³
TEMPERATURE (exact)				
°F	Fahrenheit temperature	5/9 (after subtracting 32)	Celsius temperature	°C

*1 in = 2.54 cm (exactly). For other exact conversions and more detailed tables, see NBS Misc. Publ. 286, Units of Weights and Measures. Price $2.25. SD Catalog No. C13.10:286.

Approximate Conversions from Metric Measures

Symbol	When You Know	Multiply by	To Find	Symbol
LENGTH				
mm	millimeters	0.04	inches	in
cm	centimeters	0.4	inches	in
m	meters	3.3	feet	ft
m	meters	1.1	yards	yd
km	kilometers	0.6	miles	mi
AREA				
cm²	square centimeters	0.16	square inches	in²
m²	square meters	1.2	square yards	yd²
km²	square kilometers	0.4	square miles	mi²
ha	hectares (10,000 m²)	2.5	acres	
MASS (weight)				
g	grams	0.035	ounces	oz
kg	kilograms	2.2	pounds	lb
t	tonnes (1000 kg)	1.1	short tons	
VOLUME				
ml	milliliters	0.03	fluid ounces	fl oz
l	liters	2.1	pints	pt
l	liters	1.06	quarts	qt
l	liters	0.26	gallons	gal
m³	cubic meters	35	cubic feet	ft³
m³	cubic meters	1.3	cubic yards	yd³
TEMPERATURE (exact)				
°C	Celsius temperature	9/5 (then add 32)	Fahrenheit temperature	°F

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

LIST OF ILLUSTRATIONS

LIST OF TABLES

EXECUTIVE SUMMARY

INTRODUCTION

The greater consequences of today's marine accidents, due primarily to the increased volume of oil and hazardous materials transported, along with the increased public concern for protecting the environment have provided the impetus, both nationally and internationally, for improvements in tanker safety. Domestically, the Port and Tanker Safety Act of 1978 requires a general improvement in standards relating to the qualifications and training of crews and specifically calls for the establishment of standards relating to "qualification for licenses by use of simulation for the practice or demonstration of marine-oriented skills." The U.S. Coast Guard (USCG) and Maritime Administration (MarAd) have embarked on a jointly sponsored project to thoroughly investigate the proper use of simulators as part of the mariner training and licensing process. The long-term goal of the project is to develop an information base from which positions, decisions, and actions may be formulated to raise the licensing and qualifications of mariners.

The USCG is interested in accrediting simulator-based training programs, which would be allowed as partial credit for obtaining a license. Minimum design standards for critical simulator and training program characteristics would be one alternative means of achieving such accreditation. MarAd is interested in assisting industry to cost-effectively incorporate simulator-based training into the mariner training process, particularly at the cadet level. The identification of critical factors affecting training effectiveness would be most beneficial in the cost effective design and operation of a mariner simulator-based training facility.

The initial phase of the project (Phase 1) was directed towards the development of an effective investigative methodology, development of a comprehensive data base pertaining to training system design, and the identification of knowledge gaps for the direction of subsequent research. The investigation centered on the research literature to summarize what is known about simulators and was concerned with bounding the problem and identifying the relevant issues.

To ensure that the results of this project would be responsible to the needs of the maritime community, a Training and Licensing Working Group, consisting primarily of representatives from labor, ship operators, maritime training facilities, MarAd, and USCG was formed. This Working Group has provided invaluable assistance in suggesting direction and reviewing the efforts of the research team. As a result, the Phase 1 final report has been well received by the maritime community.

During the Phase 1 investigations, it was found that many gaps existed in the empirical research literature. These gaps cover a wide range of issues relative to the simulator and its influence on training effectiveness and performance validity. The issues have been listed and prioritized in Volume II of the Phase 1 report. The Phase 2 program was designed to initiate a broad-based structure within which a systematic investigation of the many simulation and training variables could be conducted. A "screening process" approach was chosen as the principal investigative tool for this phase of the project, since it enables the economic investigation of a large number of variables.

DESCRIPTION OF EXPERIMENT

Experimental Objectives

The objective of the Phase 2 research was to identify simulator training system characteristics (e.g., horizontal field of view), the different levels of which have a substantial impact on the effectiveness of training. An initial set of potential characteristics was selected, prioritized on the basis of expected cost impact and training impact. The experiment was designed to refine this set by identifying those characteristics with the greatest impact, to enable through investigation of the most important characteristics. This initial experiment of the screening process is the first of several iterative, fractional factorial designs used to identify important factors, not to exhaustively investigate the selected characteristics. Once the important characteristics have been identified, the same data can be used with additional data obtained in subsequent phases of the research program to complete an exhaustive investigation of the most important characteristics, describing the

effectiveness of mariner simulator-based training for the variables considered. Screening designs are employed to help the investigator decide which factors merit investigation in greater detail at the next stage of the program.

Experimental Variables

Seven training system characteristics were investigated as the experimental variables for the Phase 2 initiation of the screening process methodology. These variables were chosen principally to evaluate the effect of high cost simulator and training program characteristics. The cost of simulator and training program characteristics was of primary concern since the cost effectiveness of training simulators is paramount to their successful implementation. The levels of these variables were chosen based on the following considerations: (1) representative of high and low cost extremes, (2) presently employed world-wide and reported effective for training, and (3) amenable to implementation at CAORF with little modification of existing systems.

The following simulator/training program variables and their associated levels were investigated during the Phase 2 experiment:

- Color visual scene: full color versus black and white

- Time of day: daylight versus night

- Horizontal field of view: 240 degrees versus 120 degrees

- Target controllability: independently maneuverable versus canned target ship

- Feedback methodology: augmented versus nonaugmented

- Instructor difference: instructor A versus instructor B

- Bridge configuration: full CAORF bridge versus reduced bridge

Experimental Design

To evaluate the effect of these experimental variables on the effectiveness of simulator-based training, an experimental simulator-based training program was developed. This experimental training program was administered to nine classes consisting of six chief mates each. The CAORF simulator was configured to represent a different combination of the above variables/characteristics for each group. The impact of each variable was determined on the basis of the change in shiphandling performance before and after the training program.

DESCRIPTION OF TRAINING PROGRAM

Goal

The training program for this experiment was designed to provide chief mates who were upgrading to master, with the necessary training both in the classroom and on the simulator to ensure that upon completion of the program each trainee would be able to safely maneuver an 80,000-dwt tanker in Port XYZ, under a variety of conditions. The National Transportation Safety Board (Annual Report 1977) and the Training and Licensing Project Working Group generally endorse improvements in master understanding of restricted water shiphandling. This should not be construed as the master supplanting the pilot's function, but as improved training for the master to enable him to better exercise his responsibility for the vessel by more closely monitoring or supporting the pilot's actions.

Subjects

This program was developed to specifically train chief mates who met the following requirements:

- Held a license as a chief mate of ocean or coastwise steam or motor vessel

- Had sailed as a chief mate within the past year

- Had sailed primarily on container ships

Schedule

The training was accomplished over a 9-week period. One experimental group, consisting of six students, was trained each week. On Monday all students attended a 2-hour CAORF familiarization session and were then separately administered a 45-minute pretest to establish their entry level proficiency. The actual training program was conducted on Tuesday, Wednesday, and Thursday. On Friday each student was separately given a 45-minute posttest to evaluate training gain.

Test Scenarios

On Monday each student to be trained during the week was separately administered a test scenario (pretest) to establish his entry level of skill. The administration of this test scenario required approximately 45 minutes per student. The gaming area for the test scenario was the hypothetical Port XYZ which was divided into four sequential geographic segments. The test scenario was devised to incorporate situations that were representative of each training objective (i.e., integrated/emergency shiphandling).

At the end of the training week (Friday), each student was again administered the test scenario (posttest). The results of the posttest for a particular experimental group (six students) were then compared to the results of the pretest to determine that group's training gain. This provided the primary means for establishing the relative impact of the investigated characteristics on training effectiveness. In addition, consideration was also given to occurrence/avoidance of collision and groundings, track deviation and deviation from recommended heading to determine training effectiveness.

CONCLUSIONS

The objective of this research was to determine the priority ranking of the variables/characteristics investigated. By far the most important factor affecting training effectiveness in this study was the difference in individual instructors. This suggests that instructor selection and training are crucial to simulator training effectiveness.

Very little evidence was found of important effects due to simulator design features. Where such effects were found, they sometimes favored the less expensive features. For example, a 120-degree field of view was sometimes more favorable than a 240-degree field of view. The project team feels that the principle involved is that of focusing the trainee's attention on the parts of the scene containing the information he is supposed to use rather than distracting him with extraneous visual cues. If this is the case, then the narrower field of view is preferable for initial training in problems where the relevant cues are ahead. More advanced training (to overcome distraction) or problems with key cues at broader bearings would favor a broader horizontal field of view.

One design finding that appears very clear is that night-only training is definitely not as effective as day training for problem-solving under daylight conditions. We suspect that the converse is also true, based on an inference that day visual navigation and night visual navigation rely on distinct sets of perceptual skills.

The most interesting finding is that simulator training was effective under all design feature conditions tested. Thus, even a black and white, night-only, 120 view simulator having no plan view feedback capability and only a rigidly preprogrammed traffic vessel, can provide effective short course training for some scenarios, given an effective instructor. Similarly, even a small, fairly crude bridge with televisions mounted in windows rather than a perceptually accurate projected visual display could offer some effective training.

Certain cautions should be noted regarding these conclusions. The subjects in this experiment were highly self-motivated, experienced professional mariners. The experimental design was exploratory, not conclusive. Greater rather than lesser design features can be justified for certain scenarios and for extended training programs. Finally, it is conceivable that failure to demonstrate significantly increased training value in some cases was due to insensitivity of the scenarios or performance measures rather than a genuine lack of any significant added value in the more expensive simulator features.

Variable Ranking

The methodical research approach followed in this experiment was a "screening process." It is an iterative approach allowing the obtainment of an overview of the effects of a great many variables in order to select the most important to study more precisely later in the process. Table A summarizes the resultant ranking of the variables examined

TABLE A. VARIABLE RANKING FOR
FUTURE RESEARCH
(DECREASING PRIORITY CATEGORIES)

Priority I
Instructor
Horizontal Field of View
Feedback Methodology
Priority II
Time of Day
Target Maneuverability
Color Visual Scene
Priority III
Bridge Configuration

in this experiment. They have been grouped into three categories for future research based on the observed criticality of their impact on the effectiveness of simulator-based training, as administered during this experiment. The three characteristics/variables in the priority I category should be thoroughly investigated. Those in categories II and III should receive secondary emphasis.

Simulator Design Characteristics

Research could not possibly provide the quantity of objective information required to completely design bridge simulator/training devices. The number of relevant issues, as noted in the Phase 1 report (Hammell, et al., 1980) is massive. Hence, the design of simulators must remain based upon highly structured subjective analysis. In particular, it is recommended that a highly structured training analysis approach be undertaken in the design of all simulator/ training devices (e.g., task analysis, skill and knowledge requirements, training objectives, etc.). The majority of findings obtained in this experiment are consistent with conclusions that would have been reached on the basis of a structured training analysis. Objective research data regarding particular design characteristics should be continuously developed. These data should be used as the baseline from which the large number of design characteristics would be determined; the structured training analysis would extend from the limited baseline data available to the comprehensive design of a particular simulator/training device. The consistency of results in this experiment strongly supports this approach. The most cost effective simulator/training device design is likely to be the result of a multidisciplinary design team comprised of mariner experts fully understanding tasks, etc., of the deck officer, and training experts fully understanding training technology and its application.

Simulator Fidelity

The fidelity of particular simulator characteristics (e.g., field of view) may have the potential to be manipulated so as to achieve a greater level of training effectiveness. This type of manipulation, which is alluded to in the research literature but has not been supported, results in greater training leverage as compared with the higher fidelity duplication of the actual at-sea characteristics. Under certain circumstances, the lower fidelity level may actually achieve a higher training effectiveness than the higher fidelity level. The 120-degree field of view resulted in greater training effectiveness under certain situations in

this experimental training program than the 240-degree field of view; the opposite was true, also in other situations. This appears to be due to the lower fidelity level (e.g., smaller field of view) unencumbering the trainee to some extent and assisting him in focusing his attention on that information which is most relevant. In the case of this experiment, the smaller field of view assisted the trainee in focusing his attention ahead of ownship where the relevant information existed in those particular situations and removed much of the distracting information abeam of ownship. Hence, the level of fidelity of simulator characteristics could be manipulated during the training program to achieve a more rapid training gain. Careful consideration should be given to the other aspects of the training system when enhancing the effectiveness of the training process via the use of lower fidelity characteristics. It is expected that some additional training under conditions of higher fidelity would be necessary for the trainee to adapt to the real-world situation after having acquired the basic expertise under more limited conditions. This finding coincides with the generally accepted progressive approach to training (e.g., one must crawl before he can walk).

Training Assistance Technology

The utilization of advanced concepts of training technology may have potential to greatly enhance the effectiveness of the training process. The strong effect of the instructor observed in this experiment indicates that techniques and tools which improve the instructor's ability to teach the desired skills and knowledge could have a substantial positive impact on the effectiveness of mariner simulator-based training. These may include advanced training methodologies, appropriate classroom/simulator mix tailored to the objectives, use of classroom feedback aids through curriculum materials, instructor training, amount and type of feedback, exercise design, use of performance indicators, and others. It is essential that the cost effective training system be based on careful consideration of the nonsimulator as well as simulator aspects.

Training Effectiveness

The 3-day simulator-based training program developed for this research was found to be an effective means of improving the integrated and emergency shiphandling skills of deck officers. This conclusion does not mean to imply that simulator-based training programs should be configured for 3 days in length. Rather, it shows that a 3-day training

program can be effective if properly structured. Of greater importance, it indicates the potential effectiveness of simulator-based training for deck officers. The transfer of mariner simulator-based training to the at-sea environment, although assumed, still remains to be established.

Port XYZ

The simulated Port XYZ appears to be an acceptable test scenario for future research and training at the master/chief mate level. This research establishes that an effective training program for masters could be designed and implemented within the geographic area contained in Port XYZ. In addition, it should be noted that Port XYZ contains many of the attributes recommended by the SNAME Panel H-10 (Controllability) to be included in a standardized test for evaluating the handling of large tankers entering or departing a representative port (SNAME Panel H-10, 1975).

Performance Measures

There appears to be no single performance measure for evaluating shiphandling performance in restricted water scenarios as represented by Port XYZ. The most effective method of evaluating the performance of trainees appears to be through utilization and application of several performance measures. It is expected that a simulator-based training facility with on-going training would be in a position to refine the performance measures utilized to evaluate students on their particular scenarios.

It should be noted that performance evaluation is not considered as a necessary part of the training process. Performance evaluation is necessary to determine the amount of training gain achieved over the training program, and for other evaluations of the trainee population. However, it should not be viewed as a necessary part of an effective training process. The performance measures are necessary for training, but only for the presentation of relevant performance related information to the trainees so as to enable them to learn and understand the relationships between the various parameters they are involved with (e.g., impact of rudder angle on turning circle; impact of range of maneuver on resultant CPA). The performance measures, which are extremely important for training, do not have to provide a single overall indication of performance; furthermore, a particular performance measure may often conflict with other performance measures, as is the case when the deck officer is faced with mixed results for any particular choice of actions. The performance measures, even when conflicting, will provide the trainees with

useful information to relate their actions to particular outcomes in the operational situation. Hence, a variety of performance measures should be used during any training exercise to generate and provide the necessary information to the trainees and thus enable them to learn the relationships of interest and achieve the appropriate skill levels.

RECOMMENDATIONS

The following recommendations concerning the proper design of a mariner simulator-based training system are made on the results of this Phase 2 research. Recommendations relating to future research are contained in Section 4.2.2.

- Considerable care should be given to the selection of the instructor. This research indicates that the instructor is an extremely important factor relating to the effectiveness of a simulator-based training program. Additional research is required to define the attributed of a well-qualified instructor.

- The simulator horizontal field of view should be specially tailored to the training objectives and other training situation characteristics. A variable field of view is recommended. This research strongly indicates that a reduced field of view is preferable for training selected skills as discussed in Section 3.8. Caution, however, should be exercised in manipulating the horizontal field of view so as to achieve enhanced training effectiveness. It is recommended that when a reduced field of view is used to focus attention and otherwise enhance the training process, it should be followed by some amount of training under conditions of the higher fidelity level so as to enable adequate transfer to the operational situation.

- Feedback techniques expanded beyond those employed during this research should be considered for incorporation into ship bridge/shiphandling simulators. Techniques which improve the instructor's ability to teach the desired skills could have a substantial positive impact on the effectiveness of simulator-based training. Caution should be exercised in matching the feedback techniques employed to the training objectives as discussed in Section 3.6.

- It is recommended that simulator-based training occur under the ambient lighting conditions that the skills will be used in the at-sea environment (i.e., day or

night). Although the results of this research do not firmly establish a basis for this recommendation as discussed in Section 3.7, they do provide an indication that daytime training is preferable for daytime shiphandling. In addition, lacking any information to the contrary, it appears only prudent to train on the simulator under the operational lighting conditions anticipated at-sea.

- This research indicates that the type of traffic vessel control may have a significant impact on the effectiveness of simulator-based training. The results of this experiment as regards the impact of canned versus independently maneuverable traffic vessels were inconclusive although the experimenter's interpretation discussed in Section 3.4.3 favors independently maneuverable traffic. There are undoubtedly specific training objectives where the proper use of canned traffic vessels can provide the most cost-effective training vehicle. However, until such guidelines are better defined, it is recommended that an independent traffic vessel capability be considered for any simulator involving interactions between two vessels since this capability is able to more accurately model the wide range of behavior encountered at sea.

- This research indicated very little differences in training effectiveness between a color visual scene and a black and white visual scene. This finding should not be interpreted to mean that a color visual scene is not required for shiphandling training. This screening experiment evaluated only four specific shiphandling situations within the test scenario. As discussed in Section 3.5, these situations may not have been sensitive to the use of color although they contained some of the traditional color cues used by the maritime community (e.g., traffic vessel lights and aid to navigation markings). In other situations, a color visual scene may be more critical for effective training. As a result, it is recommended when designing a simulator-based training system for senior mariners that the color cue requirements of the desired training objectives be carefully analyzed. For certain applications, a black and white visual scene may suffice. However, it is recommended that a simulator-based training facility which offers or plans to offer a comprehensive training program to have a visual scene capable of simulating color for at least vessel sidelights and aids to navigation—these being the principal color cues historically used by the maritime community.

- Training system acceptance criteria should be based on either (1) minimum training system design standards (i.e., including the simulator/training device and training program) or (2) test performance standards (i.e., periodic tests of the school's training effectiveness). It is recommended that the quantity and quality of information currently available pertaining to minimum design standards are inadequate for basing the training system acceptance criteria today; hence, it is recommended that the criteria be based on test performance with eventual evolution to a combination of test performance and minimum design standards.

SECTION 1

INTRODUCTION

1.1 BACKGROUND

Mariner training has become the focus of national and international concern over the past several years. The Intergovernmental Maritime Consultative Organization (IMCO) Convention on Standards of Training and Watchkeeping (IMCO, 1978) calls for the upgrading of mariner skills via improvement in mariner training. Several aspects of this convention, which are currently being considered for ratification by governments of member nations, address improvement in deck officer knowledge and skills from the cadet/3rd mate up through the master. Furthermore, this convention recognizes the use of the ship bridge simulator for transition training of deck officers, particularly those from underdeveloped countries which do not have the capabilities to otherwise provide needed training/experience.

In addition to the improved training standards for deck officers, the nations attending the IMCO Convention reached an agreement on substantially altering the training requirements for maritime academy cadets pursuant to obtaining the 3rd mate's license. A consideration is currently being made for substituting some simulator-based training of cadets with some at-sea training requirements. The ship bridge/shiphandling simulator appears to have good cost effective training potential to fulfill this need. The areas of application, the cost effectiveness of such training, and the necessary simulator-based training system design characteristics must be determined.

The U.S. Port and Tanker Safety Act of 1978, which follows from the Presidential Address to Congress on March 17, 1977, requires establishment of standards pertaining to the training of deck officers. In particular, standards are required with regard to vessel handling and navigation under normal and emergency situations. The establishment and/or upgrading of standards relates to a variety of training applications, such as the transition training of a deck officer going from one type of vessel to a substantially different type of vessel. Furthermore, standards are required to establish qualification for deck officer licenses by the use of simulators.

The role of the simulator, and more properly the simulator-based training system, in the mariner training and licensing process is an issue of central concern. The current availability of adequate simulation technology has brought about both the potential for highly effective deck officer training, and the requirement to substantially improve the current level of deck officer skill via training. Well over a dozen major ship bridge/shiphandling training simulators have been put into operation during the past 12 years. These facilities all have similar goals (i.e., deck officer training) although their simulator-based training system characteristics differ widely. The major difference lies in the visual scene characteristics. This aspect of the ship bridge/shiphandling simulator has proven to be the most challenging, with unique solutions developed to achieve the most cost effective designs. The visual scene characteristics of these simulators differ widely (e.g., color visual scene versus black and white visual scene) as do their respective costs (i.e., by a factor of 15 to 1). Due to this difference, it is reasonable to expect that the training effectiveness of the simulators would also differ widely with regard to particular training objectives. Unfortunately, relatively little information is currently available regarding the cost effectiveness of alternative simulator-based training system characteristics to aid the training system designer, operator, and/or user. The reason for this dearth of information is the costly and time consuming nature of training system research.

Two primary issues exist regarding the use of simulators for deck officer training: (1) the cost effective design of the simulator-based training system, and (2) the requirement of, or allowance for, simulator-based training for partial satisfaction of some licensing requirements. The use of the ship bridge/shiphandling simulator for deck officer training is gaining acceptance in the maritime industry, as evidenced by the increasing number of simulators available for such training. This trend is obviously toward the use of this technology if it can be demonstrated to be a cost effective means of acquiring and improving the deck officer skills. The need exists, therefore, to identify the cost effectiveness of alternative simulator/training system design characteristics

(e.g., horizontal field of view) in relation to the training of particular deck officer skills. In essence, considerable research needs to be conducted to assist the training system designer, operator, and user. Many issues underlie this area of investigation.

The apparent cost effective availability of simulator-based training, together with the need to improve the qualification standards for licenses, suggest that it is likely to expect that simulator-based training will be allowable for substitution of some license requirements, or required for partial satisfaction of some licensing requirements. In either case, the effectiveness of simulator-based training for meeting certain licensing requirements must be determined. In addition, it will be necessary to ensure that each simulator-based training program meets those particular licensing requirements. Hence, the need exists to determine licensing requirements for which simulator-based training can be substituted and/or should be required, and to develop a means of evaluating simulator-based training programs as meeting some minimal acceptance criteria.

The U.S. Maritime Administration and the U.S. Coast Guard have jointly embarked on a multiyear program to investigate the role of the ship bridge/shiphandling simulator in the mariner training and licensing process. This report addresses the second phase masters-level empirical investigation of this project.

1.2 LONG-TERM GOALS

The long-term goal of the project is to develop an information base from which positions, decisions, and actions may be formulated to raise the licensing and qualification standards of mariners. More specifically, the project will thoroughly investigate the potential role of simulators, and develop the information base from which recommendations may be made, in support of improving the training and licensing of mariners.

1.3 PHASE 1

The initial phase of the project (Hammell, Williams, Grasso, and Evans; 1980) was directed toward the development of an effective investigative methodology and development of a comprehensive data base pertaining to deck officer behavior and training system design. It was concerned with bounding the problem and identifying the knowledge gaps for the direction of subsequent empirical research. The investigation centered on research literature to (1) summarize that which is known about simulators and training, and

(2) identify those issues that require additional research or investigation. The following Phase 1 objectives that were investigated identify the issues relevant to the use of the simulator for training.

a. Compile task analysis data to identify the tasks performed by a deck officer

b. Identify the skills and knowledge required of a qualified merchant mariner

c. Identify the factors that influence the level of skill and knowledge required

d. Delineate the potential uses of simulators to develop and demonstrate the skills

e. Identify the factors that effect the feasibility of using simulators to improve and/or demonstrate mariner skills, relative to other means of skill acquisition

f. Delineate the issues that require further investigation

1.4 PHASE 2

As a result of the Phase 1 investigation, it was found that many specific gaps existed in the empirical research literature, particularly regarding the use of simulators for mariner training. The gaps cover a wide range of variables relative to simulation and its influence upon training effectiveness and performance validity. Phase 2 seeks to clarify several of the more important issues facing the design and use of simulators for training in the maritime industry by developing a broad-based structure within which a systematic investigation of the many simulation and training variables could be conducted. Several of the research and development areas that require further investigation, as noted in Phase 1, include:

a. Simulator fidelity
   * Color versus black and white
   * Horizontal field of view
   * Visual detail/resolution
   * Equations of motion

b. Training Assistance Technology
   * Feedback methodologies
   * Feedback displays
   * Diagnostic evaluation and placement

c. Simulator training effectiveness
   - Skills amenable to simulator training
   - Skills versus simulator characteristics
   - Types of training—upgrading, transition, refresher
   - Amount of simulator time

d. Training methodology
   - Training techniques
   - Mix of classroom and simulator
   - Use of simulator—e.g., demonstration or hands-on
   - Time compression
   - Team versus individual approach

e. Fundamental skill elements of abilities
   - Identify fundamental skills
   - Relate parameter/characteristic to vessel type
   - Identify operational cues and techniques
   - Training program construction

f. Performance evaluation tool
   - Test development
   - Integrate with the training program
   - Validation

g. Training system acceptance criteria
   - Criteria development
   - Evaluation of alternative designs

h. Curriculum guidelines
   - Content—SFOs, knowledge, and skills
   - Program structure—modular versus standard; diagnostic placement
   - Prototype course evaluation

i. Performance measures and standards
   - Performance measures development
   - Validation
   - Determine the reliability of performance measures
   - Set standards and evaluate

Phase 2 investigated several high priority issues from the above list. The selection of the issues for investigation and the establishment of goals for Phase 2 were accomplished in conjunction with the Coast Guard and Maritime Administration sponsors along with the Training and Licensing Project Working Group. This Working Group, consisting of representatives from labor, ship operators, and maritime training facilities, was formed during Phase 1 to ensure that the results of this project would be responsive to the needs of the maritime community.

The goals of Phase 2 were as follows:

a. Identify several simulator and training program characteristics that have potential impact on training system design and training system acceptance criteria.

b. Investigate the effectiveness of a simulator in training selected skills as well as the effectiveness of alternative simulator and training program characteristics in achieving an effective training process.

c. Delineate the training methodology to be employed.

d. Construct an empirical training program and implement on the Computer Aided Operations Research Facility (CAORF) to investigate the training system variables.

e. Develop performance measures to evaluate trainee performance and training program effectiveness, and hence to evaluate the issues investigated.

An empirical investigation was conducted on the simulator at CAORF, Kings Point, New York to accomplish these goals.

SECTION 2

PHASE 2: TECHNICAL APPROACH

2.1 OVERVIEW

2.1.1 GENERAL. The Phase 2 technical approach was segmented into five parts, as shown in Figure 1. The experiment and supporting materials were developed in Parts I through III, culminating in the Pre-Simulation Report. The experiment was run, data collected and analyzed in Parts IV and V.

The Part I effort, the Experimental Design, structured the experiment, identifying and selecting the various factors to conduct an appropriate and well controlled experiment. The experimental model and procedures were developed, completing the blueprint for the subsequent development of material and conduct of the experiment. The tasks in Part I drew heavily on the work accomplished during Phase 1.

The training program was developed during Part II. This included the design and development of curriculum, both classroom and simulator-based, and the specification of supporting material requirements. The supporting material was developed in Part III. The material consisted primarily of (1) simulation software, such as modification of existing visual data base characteristics; (2) classroom material, such as visual aids; (3) tests, including simulator-based pretraining and posttraining tests; and (4) training technology support, such as feedback information. The Pre-Simulation Report submitted at the end of Part III detailed the experiment, procedures, and supporting material.

Several preexperimental evaluations were conducted during Part IV. These evaluations investigated various aspects of the experiment, the training program, and the supporting material. Modifications, prior to the experiment, were made on the basis of these evaluations. Part V followed these pre-experimental investigations. It consisted of the experimental data collection, analysis, and development of the final report.

Several constraints existed regarding the design and conduct of the Phase 2 research. First, training research is inherently time consuming and expensive due to the necessity of conducting a training program under each set of selected treatment conditions. Second, the training program for each group must be evaluated through the use of a post-training test. Third, to control for variability in student input characteristics, a pretraining test is normally administered to establish the student's entry level proficiency. The change in performance from pretest to posttest is then the measure used to determine training gain for the specific set of treatment conditions. Fourth, since students have variable skill acquisition rates, randomly selected experimental groups of students are utilized in lieu of one student per treatment condition in order to statistically control for the effect of this variable.

Since the Phase 2 effort was primarily a simulator-based experiment investigating the effectiveness of various training-related variables, a multidisciplinary approach was called for, emphasizing (1) training, (2) simulator design, (3) ship maneuvering skill, and (4) experimental design and analysis. Hence, the project team consisted of a mix of disciplines including (1) training specialist, (2) simulation specialist, (3) data analysis specialist, (4) maritime consultant, (5) experimental design specialist, and (6) CAORF operation personnel.

Six simulator design variables were chosen from those investigated in a "screening process" detailed in the following sections.

2.1.2 SCREENING PROCESS. Several methodological training research approaches may be followed to accomplish the goals outlined in the introduction of this report. Appendix F discusses the alternative methodologies considered for this experiment.

The methodological approach chosen involving the use of fractional factorial designs to assess the magnitude of effect of the variables under investigation is accurately described as a "screening process." This process represents the principal investigative tool for this phase of the project. The initial stage of this process involves the selection of
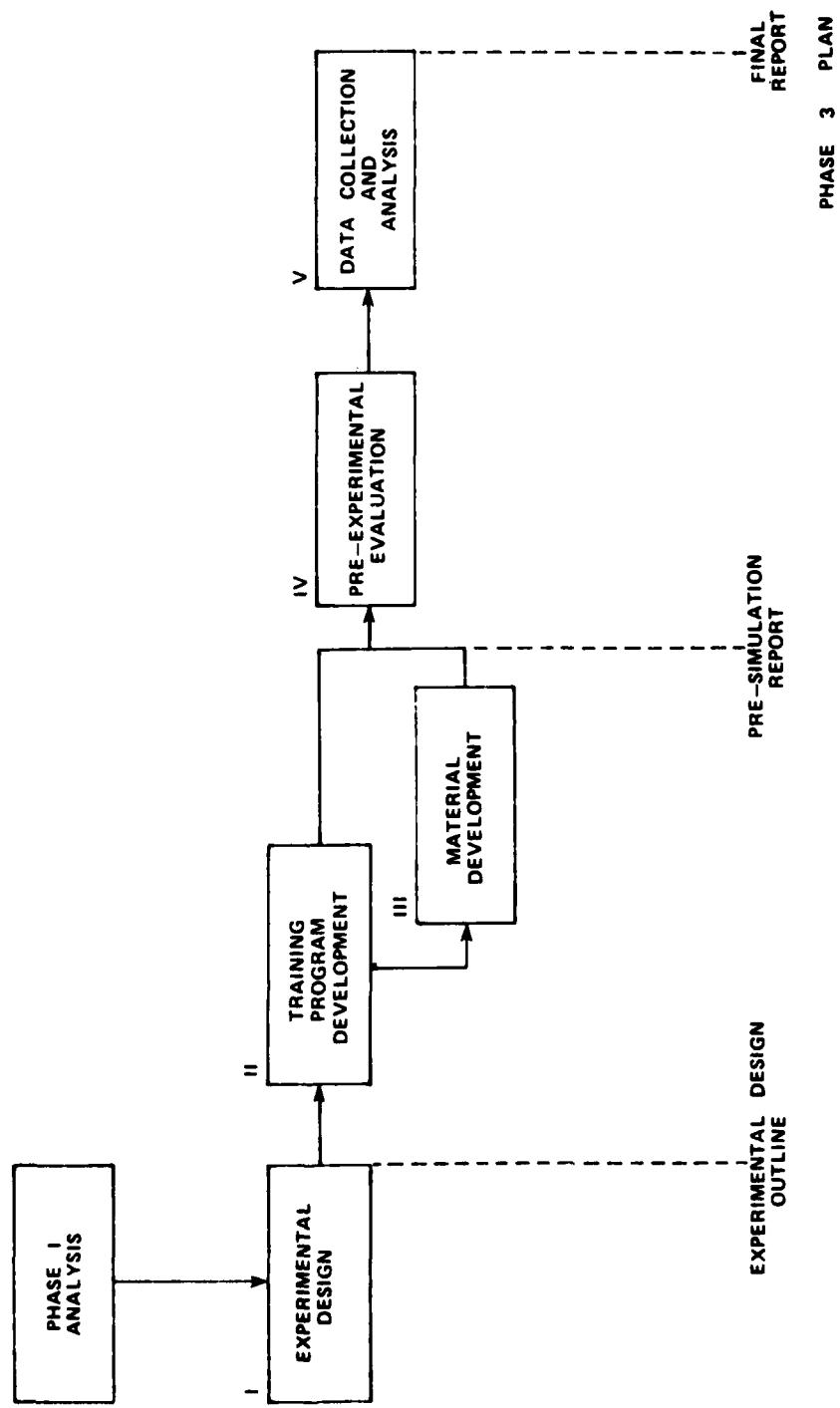
Figure 1. Phase 2 Flow Diagram

2-2

six or more variables, each at two levels and integrated into a fractional factorial design. Subjects are then run across conditions and the data analyzed by techniques such as the standard analysis of variance (ANOVA). The objective is to yield a measure of the magnitude of effect of each variable under consideration. The proportion of total variance that each variable accounts for is based on the magnitude of the effect of each variable. Once the variance of each variable is determined, those variables that contribute minimally to the variance in behavior are screened out. Those variables which provide the greatest contribution toward the variance in behavior, however, are selected for more refined analysis further along in the screening process.

Following this first stage, another set of variables would be selected and consolidated into a second fractional factorial design, adding to the data collected during the first stage. The stages are then iterated until that point when all the relevant variables have been screened to determine the proportion of variance for which each is responsible in terms of the measure of training effectiveness. The magnitude of variance of all variables investigated can be assessed after several iterations of this process. The final complement of variables would be those which have the greatest effect. The screening process eliminates the tedious process of testing all combinations of variables as would be the case if a completely balanced design was used to test the effectiveness of training related variables. Consequently, less experimental time is required to eliminate variables which have a nonsignificant effect on the behavior in question.

When the complement of variables has been selected, a more refined analysis of these variables can be conducted by use of central composite designs (i.e., Response Surface Methodology (RSM)). RSM allows for the evaluation of levels of a variable without the need for increasing the number of treatment combinations to that which would be required for a complete factorial design (Cochran and Cox, 1957; Clark and Willigees, 1972). Data collected from the central composite design can be reduced to form a prediction equation. This equation predicts the behavioral response of the system (e.g., training effectiveness in this investigation) on the basis of a particular combination of the input variable levels (e.g., ±60 degrees field of view; color; night scene). After agreement between predicted and actual behavior is proven, the prediction equation can be used to determine the training effectiveness of any number of combinations of variables and levels therein. That is, a prior assessment of differing training situations can be determined with confidence in the absence of

empirical testing if the equation derived was empirically determined beforehand and tested with confidence. The addition of new variables to the equation, however, would necessitate empirical investigation. This screening procedure then describes the broad-based framework that embodies the methodological structure for the Training and Certification program.

2.2 DESCRIPTION OF THE EXPERIMENT

2.2.1 EXPERIMENTAL OBJECTIVES. The principal objective of this screening experiment was to identify the simulator characteristic variables from those investigated that have nontrivial effects on training effectiveness. This experiment was considered to be the first of several iterative, fractional factorial designs in the screening process used to identify important factors; it alone could not achieve a complete representation of the experimental space. Once the important factors have been identified, the same data can be used along with additional data obtained in subsequent phases of the research program to complete an accurate approximation of the response surface describing the effectiveness of mariner simulator-based training for the variables considered. Screening designs are employed to enable the investigator to decide the factors that should be investigated in greater detail at the next stage of the program. As regards this project, the Coast Guard is interested in evaluating simulator-based training programs, which would be used as partial credit for meeting some license requirements. Minimum design standards for critical simulator and training program variables would be one alternate means of achieving such evaluation. The Maritime Administration is interested in assisting industry to cost effectively incorporate simulator-based training into the mariner training process. The identification of critical simulator/training program factors affecting training effectiveness would be most beneficial in the cost effective design and operation of a mariner simulator-based training facility.

2.2.2 EXPERIMENTAL VARIABLES. Six variables were selected for the Phase 2 experiment (i.e., the initiation of the screening process methodology). These variables and their associated levels were chosen on the basis of cost and potential effectiveness differences for alternative levels (e.g., ±60 degrees field of view; ±120 degrees field of view). The cost of simulator characteristics was of primary concern since the specification of a cost effective training simulator is one of the major goals of the present program, as is the validity of simulator-based training.

As a result of the Phase 1 investigation, major cost areas were identified:

a. The visual simulation

b. The complexity of computer capability with respect to the number of targets that can be presented in a visual scenario and the capability of altering target course and speed interactivity in real time

c. The real area and its associated bridge configuration

The levels of the variables that are presented below were chosen to represent reasonable high and low cost alternatives of each characteristic. The variables also represent simulator characteristics which presently are employed worldwide and have been reported to be effective for training. However, to date there is an absence of empirical data to substantiate the claims of training effectiveness made on the part of those shiphandling simulator facilities that employ the varying simulator characteristics represented in this study. The variables selected were also amenable to implementation at the CAORF facility. Subsequently, the following variables were selected:

a. Color visual scene: full color versus black and white

b. Time of day: daylight versus night

c. Horizontal field of view: 240 degrees versus 120 degrees

d. Target controllability: independent maneuverable versus canned target ship

e. Feedback methodology: augmented versus nonaugmented

f. Bridge configuration: full CAORF bridge versus reduced/reconfigured bridge

It should be noted that the "augmented" feedback condition employed several techniques in addition to verbal discussions to appraise the students of their performance during the training scenarios. A real-time display of the vessel's geographical location within the channel was supplied to a TV monitor located at the rear of the CAORF bridge. This display was for use by the instructor and those students who were not handling the vessel during the given training scenario. In addition, a hardcopy plot of the vessel's location and heading, within the channel at two minute intervals, was provided to the student who controlled

the vessel during each training scenario upon completion of the simulation exercise. These hardcopy plots were developed on the Tektronics computerized graphic terminal presently in operation at CAORF. Figure 2 illustrates a representative track plot of leg 1 of the test scenario. The "nonaugmented" feedback experimental condition used only verbal discussion to appraise the students of their performance.

The "reduced" bridge was a wood-framed module (7 by 9 feet) with five 25-inch TV monitors mounted in its windows. These monitors provided a black and white reproduction of the CAORF visual scene. This represented a reasonable composite low fidelity bridge. The following equipment was provided in the reduced bridge:

- Raytheon 16-inch radar display
- RPM order indicator
- RPM indicator·
- Speed indicator
- Wind direction indicator
- Wind speed indicator
- Heading indicator
- Rudder order indicator
- Rudder angle indicator
- Rate of turn indicator

Communications for the "reduced" bridge were provided in the following manner: The test subject was located on the "reduced" bridge. A mate was stationed at the main engine control panel on the full CAORF bridge. A helmsman was also positioned on the full CAORF bridge. Both the mate and the helmsman were instructed to carry out all orders given by the test subject via an installed intercom system.

2.2.3 EXPERIMENTAL DESIGN. Two experimental designs were employed in the Master/Chief Mate Training Experiment in order to evaluate the six experimental variables selected at two levels each. The first experimental design was a screening design used to evaluate the effect of five variables on training effectiveness (i.e., color visual, time of day, horizontal field of view, target maneuverability, and feedback methodology). To completely evaluate these experimental variables at two levels each, a 2^5 factorial design (32 conditions) would be required. In this experiment a fractional factorial design utilizing eight preselected combinations of variables was used to isolate the main effects and one two-factor interaction from another (Table 1). Each main effect, however, was aliased with strings of two-factor and three-factor interactions which will be discussed in Appendix I, Analysis Techniques.

2-4

SHIPS DYNAMIC PROGRAM—NMRC/KINGS POINT

PBT 5377   RUN NO. 1   HARBOR: PORT XYZ

START TIME 10: 0:12   STOP TIME 10:12:12   TIME INTERVAL   0: 2: 0

Vessel

Channel Boundaries

Channel Centerline
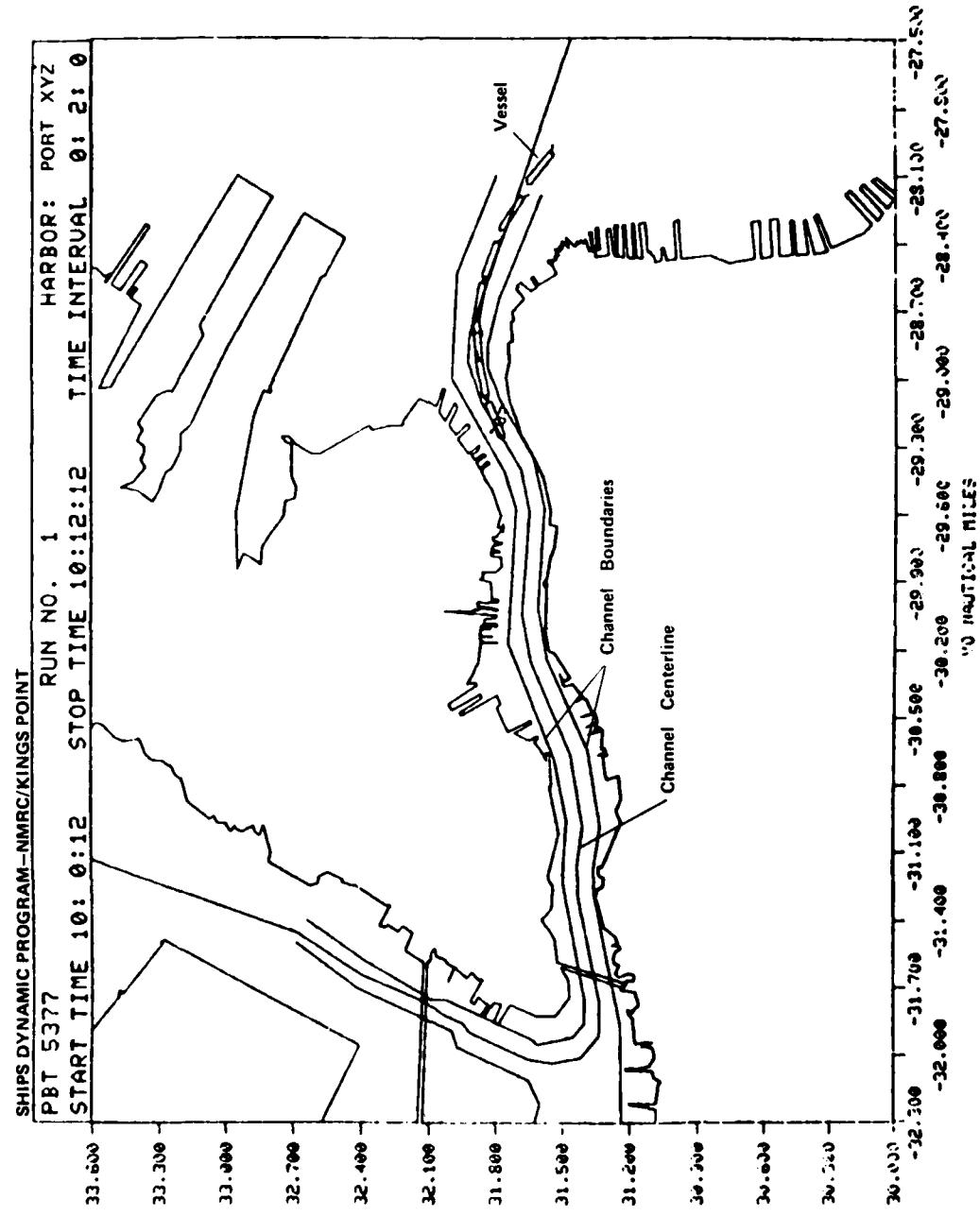
Figure 2. Track Plot — Leg 1

2-5

TABLE 1. TRAINING AND CERTIFICATION PROGRAM EXPERIMENTAL DESIGN "A"

	EXPERIMENTAL VARIABLES											
	A		B		C		D		E		F	
	Target Maneuverability		Color Visual Scene		Type of Feedback		Time of Day		Field of View		Instructor	
Groups	Independent	Canned	Full Color	Black & White	Augmented	Nonaugmented	Day	Night	120°	240°	A	B
1		•		•		•		•	•			•
2	•			•	•			•	•		•	
3	•			•		•	•			•		•
4		•	•		•			•		•		•
5		•	•			•	•		•		•	
6	•		•		•		•		•			•
7	•		•			•		•		•	•	
8		•		•	•		•			•	•	

The experimenter must determine the wisdom of neglecting these interactions, or investigating them, as he proceeds further with the screening process. The contribution of these individual effects to the total variance was then evaluated using statistical techniques (e.g., analysis of variance (ANOVA)).

It should be noted that in this experimental design, due to logistical constraints, two instructors were required to train the eight experimental groups. In order to control for any variance due to the instructor, the instructor was incorporated as a variable into the experimental design. Instructor "A" trained groups 2, 5, 7, and 8 while instructor "B" trained groups 1, 3, 4, and 6.

The second experimental design employed was a comparative design used to evaluate the relative effectiveness of the reduced bridge versus the full bridge (i.e., bridge configuration). A ninth group was trained on the reduced bridge under treatment conditions identical to group 3, thus allowing for a direct comparison between the performance of group 3 on the full bridge and that of group 9 on the reduced bridge (Table 2). This variable was not included in the screening process design along with the other variables because of its unique nature. The bridge configuration is not a "pure" variable but a composite of many variables, such as bridge size, console configuration, and type of visual presentation, designed as one. To have included it in the screening design would have confounded the analysis of the other variables.

The performance score (dependent variable) for each unit of both designs represented the change in behavior for that treatment condition. This score, which is based on a sample size of six students for each unit, is calculated as the pretest/posttest difference on the test scenario. The test scenario was given prior to and upon completion of the training program.

TABLE 2. TRAINING AND CERTIFICATION PROGRAM EXPERIMENTAL DESIGN "B"

	EXPERIMENTAL VARIABLES											
	A		B		C		D		E		F	
	Target Maneuverability		Color Visual Scene		Type of Feedback		Time of Day		Field of View		Reduced Bridge	
Groups	Independent	Canned	Full Color	Black & White	Augmented	Nonaugmented	Day	Night	120°	240°	A	B
3	•			•		•	•			•	•	
9	•			•		•	•			•		•

2.3 DESCRIPTION OF TRAINING PROGRAM

2.3.1 GOAL. The training program for this experiment was designed to provide chief mates upgrading to master with the necessary training both in the classroom and on the simulator, to ensure that each trainee upon completion of the program would be able to safely maneuver an 80,000 dwt tanker in Port XYZ, under a variety of conditions. The National Transportation Safety Board (Annual Report 1977) and the Training and Licensing Project Working Group both endorse improvements in master understanding of restricted water shiphandling. This should not be construed as the master supplanting the pilot's function, but as improved training for the master to enable him to better exercise his responsibility for the vessel by more closely following the pilot's actions.

2.3.2 TRAINING OBJECTIVES. Phase 1 of the Training and Licensing Project identified a set of 74 specific functional objectives (SFOs) which represent those deck officer skills to be achieved in restricted waters. These SFOs were grouped into five major categories. These categories were:

a. Fundamental shiphandling: Objectives that require the understanding of how a vessel will respond based on variables such as the vessel's configuration, its mass, its power (or lack of it), its reaction to currents, winds, interactions, speeds, and control actions. These variables, though constantly changing, must be resolved into definite rudder or engine orders.

b. Integrated shiphandling: Objectives which address the skills required to successfully handle the vessel in all types of situations (e.g., conning through a channel, docking, mooring, anchoring) and under various conditions, while taking into account the combination of fundamental variables which will affect the vessel's response.

c. Emergencies: Objectives which require the understanding of vessel characteristics to allow for the proper shiphandling decisions to be made and, if possible, perform corrective ship control actions to successfully ensure vessel and crew safety when personnel ship control errors or power, rudder, equipment, or electrical failures occur.

d. Team coordination/communication: Objectives which require each team member to perform parallel and serial functions in coordination with the other team members in a timely manner and within a framework of set procedures which are situation dependent.

e. Bridge procedures: Objectives which require the bridge team to organize and carry out the duties and pattern of communications required to properly execute the port entry/exit passage plans, especially when the unexpected arises.

Based on discussions with the working group and a review of current simulator training curricula, two of the five SFO categories — integrated shiphandling and emergency shiphandling — were selected for the Phase 2 investigation. As it was not feasible to train all the SFOs comprising each category, the following objectives were selected for inclusion in the training program based on (1) the appropriateness and potential to train on the simulator; (2) feasibility of training within the program constraints; and (3) dependence on the independent variables and their levels.

a. Integrated shiphandling

● Maneuver the vessel through the channel maintaining intended track when either:
  – The navigation range structures available for various channel legs have a light extinguished, one or both range structures obscured, or one structure missing.
  – The buoys available for various legs of the channel are extinguished, off position, or missing.

● Maneuver the vessel through sharp bends or blind turns into a "Y" channel maintaining intended track and safely avoiding any vessel traffic.

b. Emergency shiphandling

● Maneuver the vessel, maintaining ship control as best as possible, when a rudder failure occurs in any channel leg or turn.

● Safely maneuver the vessel when a degradation in the amount of power, or a complete power failure occurs in any channel leg or turn.

The training program was segmented into seven training units which spanned the three days of the actual training. Four training units addressed specific integrated shiphandling objectives while three training units addressed specific emergency shiphandling objectives.

2.3.3 SCHEDULE. The empirical data collection was accomplished over a 9-week period. One experimental group, consisting of six students, was trained each week. On Monday all students attended a 2-hour CAORF familiarization session, which included hands-on maneuvering, and were then separately administered a 45-minute pretest to establish their entry level proficiency. The actual training program was conducted on Tuesday, Wednesday, and Thursday. On Friday each student was separately given a 45-minute posttest. Figure 3 outlines this schedule.

2.3.4 SUBJECTS. This program was developed to specifically train chief mates who meet the following requirements:

● Hold a license as a chief mate of ocean or coastwise steam or motor vessel

● Have sailed as a chief mate within the past year

● Have sailed primarily on container ships

A test subject demographic analysis was completed and is detailed in Appendix E of this report.

2.3.5 FAMILIARIZATION. The CAORF familiarization session consisted of a 1-hour classroom session followed by a 1-hour simulator session. During the classroom session, the students were provided an introduction to CAORF and briefed on the purpose of the experiment, their schedule for the week, and the simulated geographic area where the training would be conducted. In addition, the Student Handout Package contained in Appendix A was distributed. During the simulator session, the students were instructed on the use of the hardware in the bridge, introduced to the CAORF visual scene, and allowed to handle a containership during open-water maneuvers. The latter exercise allowed the students to make the transformation from their shiphandling experience on a containership at sea to a comparative vessel in the simulated environment. The principal purpose of this procedure was to acclimate the student to the simulator prior to the pretest. In addition, it was hoped that this exercise gained the student's confidence in CAORF's ability to depict the real world and enhanced his motivation during the training program.

2.3.6 INSTRUCTIONAL METHODOLOGY. Each of the seven training program units was comprised of two segments: a classroom exercise and a simulator exercise. One-hour blocks of classroom and simulator time were desired. However, to provide for the more effective utilization of CAORF during classroom periods, the training program was restructured into 2-hour simulator blocks.

An instructor's guide was developed for the training program which specifically outlined the instructional techniques to be employed during each training exercise. Sample sheets from the instructor's guide are contained in Appendix G. The basic format for the classroom exercise was:

● Instructor presentation of scenarios within the exercise

● Seminar discussion on alternative maneuvering techniques for each scenario

● Positive guidance by the instructor on acceptable maneuvering technique for each scenario

The basic instructional format for the simulator exercise was:

● Brief review by the instructor prior to each scenario

2-8
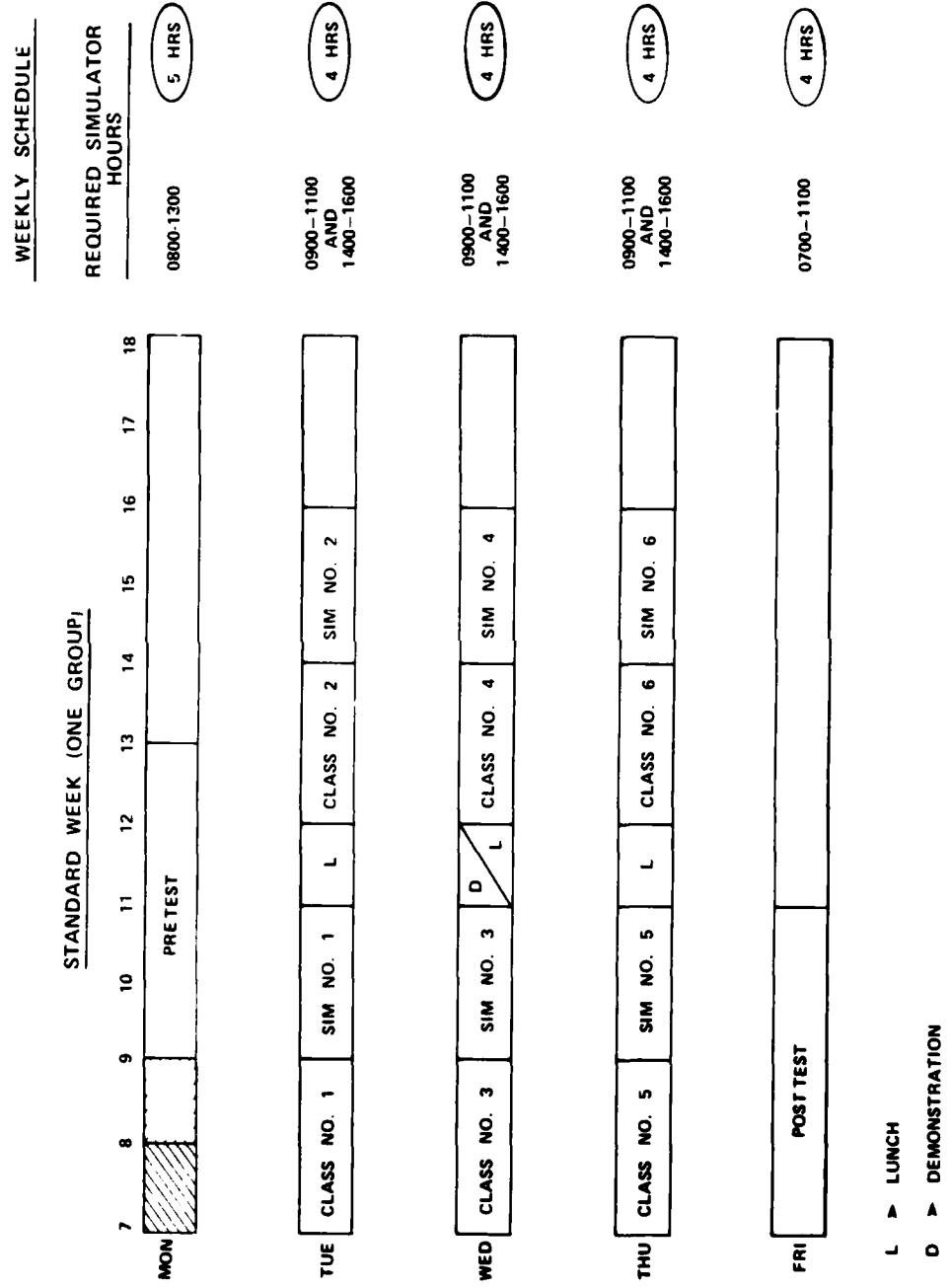
Figure 3. Experimental Group Student Schedule

- Positive guidance by the instructor during the scenario (for the groups with the augmented feedback condition, this involved the use of the situation display with those students not maneuvering the vessel)

- Postproblem critique by the instructor after each scenario

A comprehensive set of visual aids for classroom use was developed. These visual aids were used by the instructor to illustrate the results of various strategies employed by ownship when faced with various integrated or emergency shiphandling situations.

2.3.7 INSTRUCTORS. Instructors utilized for the master/chief mate training program possessed qualifications that included the following characteristics:

- Master's license for ocean or coastwise steam or motor vessel

- Experienced shiphandler

- Familiarization with the CAORF simulator

- Instructor experience

The instructors for the master/chief mate training program were briefed on the objectives of the experiment, the instructional techniques to be employed, the associated training scenarios, and recommended procedures for maneuvering the vessel prior to the commencement of the training program.

2.3.8 INSTRUCTIONAL MATERIALS. A comprehensive set of visual aids for classroom use was developed. These visual aids included (1) slides of the visual scene, (2) vector diagrams of the wind and current acting on the vessel, and (3) trackplots of vessel's response after selected rudder and engine orders were implemented. The latter were developed using an off-line computer model. Samples of these visual aids are contained in Appendix H. These aids were used by the instructor to illustrate the results of various strategies employed by ownship when faced with various integrated or emergency shiphandling situations.

2.3.9 TRAINING SCENARIOS. A total of 30 scenarios were developed for utilization throughout the 3-day training program. Each of these scenarios was constructed to achieve a particular training objective and to be sensitive to one or more of the experimental variables. For example,

one integrated shiphandling scenario in which the vessel is maneuvered through a 129-degree turn marked by several visual aids to navigation (buoys and beacons) with 1.5 knots of ebb current was observed to be sensitive to the following variables: field of view, day versus night, color versus black and white, and type of feedback.

All training scenarios occurred in a hypothetical Port XYZ which is illustrated in Figure 4. The dominant characteristics of Port XYZ include:

- 4.5 nm, 800 to 1200-foot wide dredged channel

- 129-degree turn

- 59-degree turn

- Junction

- Bifurcation

- Assortment of aids to navigation including one range

- 5-foot under keel clearance

- 3.0-knot maximum current

- 50-knot maximum wind

- 8 knots — vessel speed restriction

- No traffic control

- Moderate tug and barge traffic

It should be noted that each student observed all 30 training scenarios. Each student had hands-on simulator experience for a minimum of six training scenarios distributed across the training objectives. Table 3 outlines the relationships between training scenarios, simulator exercises, and student hands-on experience for this training program. It should be noted that a given simulator exercise consists of several scenarios and occupies one 2-hour block of simulator time. Several scenarios are repeated during the course of instruction as annotated by "/1," "/2," or "/3," after the scenario number.

2.3.10 TEST SCENARIOS. On Monday each student to be trained during that week was separately administered, for evaluation purposes, a test scenario (pretest) which is described in detail in Appendix B. The results of the

Figure 4. Port XYZ

TABLE 3. SIMULATOR EXERCISES

Scenario	Hands-on Subject
Simulator Exercise 1	
1	1
2	2
3/1	3
4	4
5	5
6	6
7	1
8	2
3/2	6
Simulator Exercise 2	
10	4
11	5
12	3
13/1	1
14/1	2
14/2	4
Simulator Exercise 3	
15	6
16	4
17	5
18	3
19	1
20	2
Simulator Exercise 4	
21	3
22	6
23	5
24/1	4
24/2	3
25/1	1
25/2	2
Simulator Exercise 5	
26	6
27	2
28/1	1
28/1	6
13/2	5
Simulator Exercise 6	
29/1	3
29/2	4
29/3	5
30/1	1
30/2	6
30/3	5

pretest provided an estimate of each student's baseline performance. The administration of this test scenario required approximately 45 minutes per student for completion. The gaming area for the test scenario was the hypothetical Port XYZ, which was divided into four sequential geographic segments. Each leg was developed to address a particular training objective. A description of each of the four legs is as follows:

- Leg 1 — Ownship was positioned in Wyassup Bay at the entrance to Gibson's Channel. The ownship position was misaligned on the ferry point range. Furthermore, an anchored vessel obscured both range lights. The wind velocity was 15 knots, gusting to 25 knots from the southwest. A constant 1.5 ebbing current was present. Ownship's objective was to navigate through the turn into Gibson's Channel. A tug was proceeding from Gibson's Channel into upper Wyassup Bay, creating a crossing situation with ownship.

- Leg 2 — Ownship's position was appropriately aligned in the channel west of the turn into Wyassup Bay. The initial ownship course was 270 degrees true. Wind and current remained the same as in leg 1. Ownship's objective was to pass through Gibson's Channel towards Fisherman's Point. A 4-minute rudder failure with a rudder angle of 10 degrees, occurred shortly after the beginning of leg 2.

- Leg 3 — Ownship was positioned about halfway through Gibson's Channel, appropriately aligned. Initial course was 247 degrees true. Wind and current were the same as in leg 1. A 3-minute propulsion plan failure commenced shortly after the start of leg 3; steering and the bow thruster were available.

- Leg 4 — Ownship was positioned in Gibson's Channel just east of the bridge, aligned appropriately. Ownship's heading was 260 degrees true. Wind and current were the same as in leg 1. Ownship's objective was to turn around Fisherman's Point entering Shellfish Bay. Two tugs were standing just east of Blackfish Island. A tug and tow was passing through the Shellfish Bay Channel to turn around Fisherman's Point into Gibson's Channel, providing some interference with ownship's plans in making this difficult turn.

On Friday, at the end of the training week, each student was again individually administered the test scenario (posttest). The results of the posttest for a particular experimental group (six students) were then compared to the results

of the pretest to determine that experimental group's training gain for the given treatment condition.

2.4 EXPERIMENTAL PROCEDURES

2.4.1 STUDENT FAMILIARIZATION. The procedures used to familiarize the student with CAORF, the purpose of the experiment, the objectives of the training program, and the operation of bridge equipment are described in the Familiarization section of the Training Program. In addition, demographic information was collected from each student. A sample demographic data collection sheet is contained in Appendix C. A test subject demographic analysis was completed and is detailed in Appendix E of this report.

2.4.2 PROTOTYPE EXPERIMENTAL RUNS. The training scenarios, test scenarios, and experimental variables were evaluated on the CAORF simulator prior to the experimental runs. The purpose of this evaluation was to ensure that the scenarios were (1) realistic, (2) accurately incorporated into the simulator, and (3) capable of being executed in the allotted time. In addition, selected training scenarios were viewed after the simulator was modified by each experimental variable. Appendix D describes the specific modifications to the CAORF simulator in order to obtain the selected simulator/training program treatment conditions.

2.4.3 DATA COLLECTION. During the experimental runs, data was collected in two principal forms: ship motion parameters and human factors parameters. The ship motion parameters utilized were the standard parameters used in other CAORF experiments. These included vessel position, velocity, heading, yaw rate, rudder position, propeller rpm, etc. These parameters were sampled at 12-second time intervals. The human factors parameters, which were collected manually from the human factors station by an experienced observer, are outlined on the sample data collection sheet contained in Appendix C. The student's performance on the test scenarios was also recorded on video tape.

2.4.4 STUDENT DEBRIEFING. During the nine weeks of the experiment, many students informally voiced their ideas concerning the training that they were receiving and their opinions of simulator training in general. To preserve this information, each subject individually participated in a debriefing session at the end of the week, which took the form of a structured interview. Questions were posed with a 1-through-10 weighting scale (1 being a low rating and 10 being a high rating). Additional comments concerning the topic were also recorded. Areas covered by the interview were student assessment of own performance, the training program, the instructor performance, the vessel characteristics, and the simulator/training variables. A sample debriefing form and an analysis of the response are contained in Appendix K.

2.5 DATA ANALYSIS

2.5.1 GENERAL. The first step of the data analysis was to establish the criteria that would provide the basis from which the experimenter could make the trivial/nontrivial judgement. Performance scores based on several performance measures were calculated for each training objective within the test scenario. The performance measures used in the analysis are explained later in this section of the report. These performance scores were used to calculate training effectiveness, which was then used to evaluate the experimental variables. Training effectiveness (i.e., the posttest score minus pretest score for a given performance measure) was used for this analysis in lieu of simply the posttest score. This was done to control for the variability of the entry level skills of students participating in the training program, so as not to bias the results of the analysis.

Cohen and Cohen (1975) point out the difficulty of correlating experimental treatment conditions with simple change (posttest minus pretest) due to the attenuation effect of the measurement errors associated with both pretest and posttest scores. They identify simple change as the conservative approach, protecting the research from reporting a false positive claim (i.e., that a significant difference was found when in fact the difference was merely due to sampling error). However, since this experiment was a screening design and its purpose, for future research, was to eliminate variables which have trivial effects related to training effectiveness, it was desirable to be protected against falsely eliminating any variables that did have nontrivial effects. As a result, attention was given to Cohen and Cohen's techniques of correcting for the attenuating effects of this measurement error through the use of a "B" correction factor applied to the pretest. The calculation of this "B" correction factor uses the pretest and the posttest correlation coefficient. Since analysis of the experimental data revealed negative correlations and since Cohen and Cohen's techniques have unresolved difficulties with negative correlations, we reverted to the use of simple change. It is believed that the use of multiple performance measures will provide protection from falsely eliminating variables from future research which have nontrivial effects.

2-13

The second step was to statistically analyze the training effectiveness data utilizing several analysis techniques. This was completed so as to gain an understanding of the influence each experimental variable had on the recorded change in performance. It should be noted that this process was applied for each training objective, since it was expected that the importance of the experimental variables would vary between integrated shiphandling and emergency shiphandling training objectives. The employed analysis techniques, which are discussed in Appendix I, include the following:

- Magnitude of Effects Calculation

- Analysis of Variance (ANOVA)

- Proportion of Variance Calculation

- Mann-Whitney U Test

- Fisher Exact Probability Test

- Difference of Means Tests

- Homogeneity of Variance Test

2.5.2 PERFORMANCE MEASURES. Even for fairly straightforward trackkeeping exercises (legs 2 and 3 of the present experiment), optimum characteristics of shiphandling performance measures are not easily defined. For both the waterway and ownship, reference data from which measurements are taken must be specified. With respect to geographic reference data, channel centerline is often used although it is recognized that for numerous reasons the middle of a channel is not necessarily the optimum location for ownship. An alternative "optimum track" may more realistically approximate the "best" path for transit. The optimum track can be established or defined in several ways. Each shiphandler might, by prior "passage planning," chart (either physically or verbally) the route he would intend to take for a given set of geographic, environmental (wind and current), and ship maneuverability conditions. By means of the same method, a "population" rather than "individual" equation of optimum track might be accomplished through averaging the intended tracks of a sample of shiphandlers. Or, the optimum track for a population might be estimated "ex post facto" through an averaging of "actual" rather than "planned" transits. A third version of optimum track consists of an abstract or "absolute" track based not upon subject intent or practice but rather on the instructor's

determination of best transit based upon textbook principles of shiphandling and personal experience. All reference data noted, however, share the same limitation in that the "best" track for any subject changes continuously with his displacement from previously intended goals. Shiphandlers generally do not try to immediately regain an intended track once they (for any number of reasons) stray from it; rather, a new "best" route is planned to some point further downtrack from the vessel's present location.

Counterpart to the choice of geographic reference data is the more limited selection of those characteristics of "ownship" from which performance measures (PMs) might be taken. Most basic is the choice between a "point" reference datum (such as center of gravity) and the volumetric calculation of "swept path." In that the present experiment entails a comparative rather than absolute evaluation of shiphandler's performance, the more economical "point" reference datum is employed. However, for calculation of CPA between ownship and moving or stationary traffic ships/obstructions, measurement is taken from "skin to skin" of the two vessels.

The measures themselves consist of "distances" (linear displacement from reference data), "angles" (ship's head deviation from channel heading, or angle to channel face at excursion), and "speed" (speed at excursion). Several numerical calculations are available —the mean and maximum/minimum deviation (CPA) being the two most common measures — but the calculation of the deviation at a specific critical point in a scenario may also be warranted, especially in emergency scenarios. With respect to "heading" PMs, mean calculations are appropriate only in scenarios (legs) in which wind and current conditions are such that there is no tactical advantage to "crabbing" or maintaining a heading other than the course on which the shiphandler intends to proceed. Similarly, absolute (single point) heading deviation is appropriate only when it is a function of the shiphandler's effort to minimize a maladaptive course rather than as a tactical procedure to maintain a desired course. In the present experiment such an absolute measure (deviation from channel heading) was employed in the rudder failure leg.

The numerical values of measures are based upon dimensions of magnitude, direction, or a combination of the two. How to treat the magnitude of the deviation from (or closeness to) a reference datum is also problematic. The numerical values of interval or ratio scales may not be, in fact, magnitudes of the attribute being measured (i.e.,

displacement from reference). Even when such distance units are perceived by the subject, equal increments or decrements in the physical values of the measurement are usually not given equivalent weight. That is to say, a shiphandler is not likely to weigh equally (in terms of his concern) his first 50-yard displacement from intended track with additional 50-yard increments. Or, for another example, a CPA of 100 feet is obviously better than 10 feet, but it is doubtful that an additional 100 feet of CPA would be considered to be of equal value. Therefore, in some cases the magnitude of a measure reflects this weighted characteristic, with unweighted measures implying equivalence of magnitude value on the part of shiphandlers.

Direction of deviation often requires a similar weighting. The intent of the shiphandler, conscious or not, is not always to remain as close as possible to the center of the channel. His optimum track may be influenced by geographic characteristics of the waterway, i.e., one or two-way traffic, presence of stationary objects, and variation in the slope and consistency of the channel bottom and walls. He may also favor one side over the other due to the presence of ephemeral environmental conditions such as current and wind. Another example is the tactical preference of shiphandlers to deviate to the inside rather than outside of channel centerline preceding a turn. Since shiphandlers do in fact "weight" displacement from intended track in terms of magnitude and direction, then PMs might be correspondingly weighted once again, as in choice of reference datum, by enlightened evaluation of the usual practice of shiphandlers.

All the preceding considerations (geographic and ownship reference data, characteristics of measures) apply in the case of simple trackkeeping, but are even more problematic in maneuvering scenarios in which ownship must negotiate turns, obstructions, traffic vessels, etc. In testing of integrated shiphandling skills, a further complication is introduced in this matter of PMs. Because the shiphandler is now tasked with more than trackkeeping and is required to maneuver his vessel around channel bends and/or traffic vessels, single PMs, no matter how realistic the reference datum or how carefully weighted, are inadequate. For example, deviation from channel centerline becomes less meaningful as a single PM in scenarios in which the presence of other vessels recommends a departure from the best track to be followed in the absence of traffic. Conversely, magnitude of closest point of approach (CPA) to traffic ships is less meaningful in restricted waterway scenarios unaccompanied by channel position measures that indicate how well ownship remained in the channel.

Rather than pursue the elusive goal of identifying the single PM that most closely approximates the intent of shiphandlers in transiting restricted waterways, the project has attempted to develop numerous PMs which, individually and combined, approximate the more complex standards upon which pilots base their control decisions. Two categories of PMs or standards of shiphandling behavior ("Human Factor" and "Ship Response") were employed in the evaluation of overall training effectiveness and the relative contribution toward training effectiveness of both these levels for each environmental variable.

2.5.2.1 Human Factor Performance Measures. Human factor PMs are those direct measures of shiphandler action and behavior relating (in this experiment) to shiphandling skills. Frequency of engine order, rudder order, or bow thruster usage are examples, as well as measurement of "reaction time," or time interval between the subject's recognition of an emergency and first subsequent control response. Human factor information consists either of measurement of subjects' physiological conditions or of their movements and actions during the course of an experiment.

Ship response PMs are also measures of the subjects' shiphandling skills, but entail assessment of the final result rather than initial action (human factor PMs) of ship control. Ship response PMs are not only more significant measures of shiphandling skill in this "bottom line" characteristic, they are easier to evaluate than human factor direct control measures (e.g., it is fairly obvious that the location of ownship within rather than outside the channel limits can be valued positively (a ship response measure). It is not clear that a greater or lesser frequency of rudder orders is a good or bad indication of shiphandling skill (human factor measure). Because of this evaluation difficulty with respect to human factor PMs, they are used in this study only as a supplement to the ship response data to provide more detailed information as to what actual control behavior might have contributed to shiphandling success/failure. Such a convoluted task as shiphandling in restricted waters precludes any reliance upon simplistic PMs in the evaluation of subjects' execution of test runs.

Even though performance is indeed physical and in some instances subject to precise measurement, the dependent variable (training effectiveness) is intrinsically behavioral and therefore not so easily quantified.

The following list itemizes the several human factor PMs utilized in this experiment.

1. Engine order frequency (all legs) — Number of engine orders per minute.

2. Rudder order frequency (all legs) — Number of rudder orders per minute.

3. Bow thruster order frequency (legs 2 and 3) — Number of bow thruster orders per minute.

4. Reaction time (leg 2) — Interval in seconds between notification of emergency (steering/propulsion failure) and initialization of control action (rudder/engine order). Appropriate for leg 2 only, because in leg 3 response may consist of decision not to take immediate control action.

2.5.2.2 Ship Response Performance Measures. Ship response data are computer generated based upon human factor (control device) inputs and vessel/environment dynamic response. Information of the ship's condition (position, speed, heading, course, etc.) can be obtained at time and/or distance intervals and stored on tape to be later subjected to direct statistical treatment, alphanumeric printout, or by graphic plan view. Graphic presentations of the ship's response (track plots) are valuable for simulator training purposes because of the effectiveness of a graphic review of the vessel's progress throughout a scenario, immediately following an exercise. This, in fact, was one of the subjects of examination in this experiment to determine the direction and degree of training effectiveness contingent upon the level of "feedback" (augmented versus nonaugmented). For the same reasons (visual impact, availability), track plots are valuable to analysts who may use them to refine the complete raw data analysis program to follow. In the process, a preliminary analysis of experimental conditions is accomplished although it is limited by a lesser number of data points, limited number of graphic PMs, and subject estimation of PM evaluations.

The following list itemizes the several individual and composite ship response PMs employed in the experiment and the methods by which they were obtained.

Numeric

1. Mean distance from channel centerline (all legs) — Average distance in feet of ownship's CG from channel centerline. Calculation is based upon position information for each data line from initialization to completion of leg (see Figure 5). Deviation measurements for data lines not attained due to channel excursion are assigned as constants greater than the maximum distance from channel boundary.

2. Mean distance from recommended track (legs 1 and 4) — Identical calculation to number 1 with exception of geographic reference datum which, in this case, is an optimum track recommended by the instructor. Calculations are made, from those posttest runs in both legs 1 and 4, which most closely approximate the models given in instruction. This PM was not employed in legs 2 and 3, which were essentially trackkeeping scenarios involving no maneuvering, and in which the optimum track was identical to channel centerline.

3. Distance from channel centerline at D.L. No. 40* (leg 2) — Absolute deviation (in feet) from channel centerline at data line No. 40 (Figure 5).

4. Distance from channel centerline at D.L. No. 58* (leg 3) — Absolute deviation from channel centerline at data line No. 58 (Figure 5).

5. Deviation from desired heading at D.L. No. 40 (leg 2) — Absolute deviation (in degrees) from desired heading.

6. Mean deviation from desired heading (legs 2 and 3) — Average deviation from desired heading.

7. CPA (legs 1 and 4) — Closest point of approach between ownship and all traffic vessels or vessel obstruction (tugs at anchor, moored containership) measured in feet from "skin to skin."

8. Mean distance from channel centerline and CPA (legs 1 and 4) — Composite PM weighing average distance from centerline against CPA to traffic vessels and obstructions.

9. Mean distance from recommended track and CPA (legs 1 and 4) — Composite PM weighing average distance from recommended track against CPA to traffic vessels and obstructions.

*Note: Several performance measures utilized the value of selected vessel motion parameters at distinct locations in the leg in lieu of the mean of the selected vessel motion parameters over all the data lines in the leg. It was anticipated that this would increase the sensitivity of the performance measures to change in resultant performance.
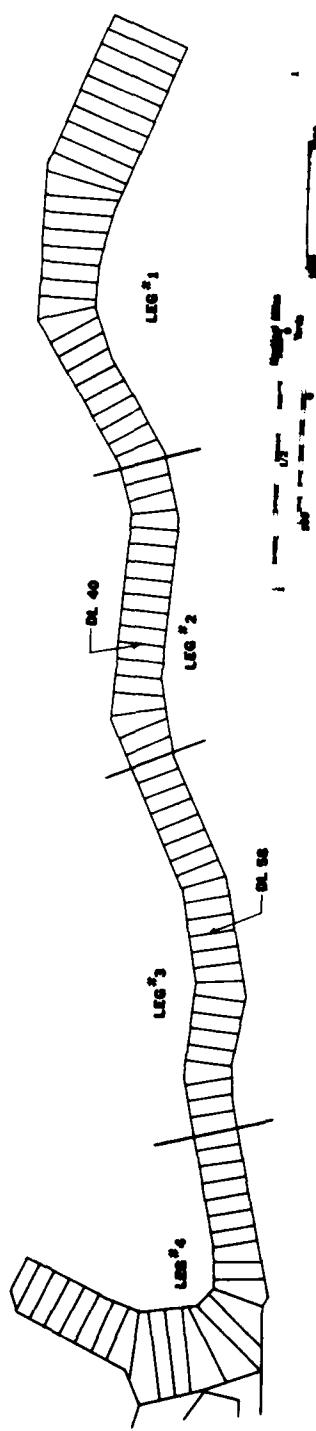
Figure 5. Location of Geographic Data Analysis Lines with Port XYZ

Graphic

10. Pass/Fail (legs 2 and 3) — A determination of whether or not ownship remained within channel limits.

11. Pass/Fail (legs 1 and 4) — A determination of whether or not ownship remained within channel limits without hitting traffic vessels or obstructions.

12. Composite (legs 1 and 4) — A composite PM is attained for legs 1 and 4 by combination of pass/fail information with measures of how well subject performed in either keeping their vessels in or leaving the channel. Relative success in remaining in the channel is designated by an average of position information (every 2 minutes) weighted according to (1) distance from channel center, and (2) direction of displacement. A 5-point scale was devised which accounts for both of these dimensions and is represented below. The directional weighting indicates the overall performance for displacement to right rather than left of center (Figure 6).

Performance in channel excursion is evaluated on the basis of (1) distance downtrack attained before excursion, and (2) the angle with the channel boundary at which the excursion takes place. In leg 1, the channel was divided into nine segments for distance measures; 1, 2, and 3 noting the sequence of outer bend sides and S, M, and L (short, medium, long) referring to gross distance distinctions for each of the three channel limit segments (Table 4). Angle of excursion was judged between 15 degrees and 60 degrees and scored with the more acute angle receiving the highest score. Comparable measures for distance downtrack and angle excursion are given in leg 4 in Tables 4 and 5.

13. Composite (leg 2) — A composite PM consisting of a determination of whether or not ownship remained within channel limits, plus a gauge of relative failure (channel excursion). The latter is evaluated on the basis of angle at excursion as weighted in Table 6.

14. Composite (leg 3) — A composite PM consisting of a determination of whether or not ownship remained within channel limits, plus a gauge of relative success (average channel position). The latter component of the combined PM is based upon the method described for No. 13.

TABLE 4. GRAPHIC PERFORMANCE MEASURE DOWNTRACK POSITION AT EXCURSION (LEGS 1 AND 4)

Leg 1	1S	1M	1L	2S	2M	2L	3S	3M	3L
	1	2	3	4	5	6	7	8	9

Leg 4	1S	1M	1L	2S	2M	2L
	1	2	3	4	5	6

TABLE 5. GRAPHIC PERFORMANCE MEASURE ANGLE OF EXCURSION (LEGS 1 AND 4)

Leg 1	15°	30°	45°	60°			
	4	3	2	1			

Leg 4	Glance	15°	30°	45°	60°	75°	90°
	7	6	5	4	3	2	1

TABLE 6. GRAPHIC PERFORMANCE MEASURE ANGLE OF EXCURSION (LEG 2)

Stern Glance	Glance	15°	30°
4	3	2	1

LEFT CHANNEL BOUNDARY

1

2

3

— CHANNEL C_L — 5 —

4

3

2
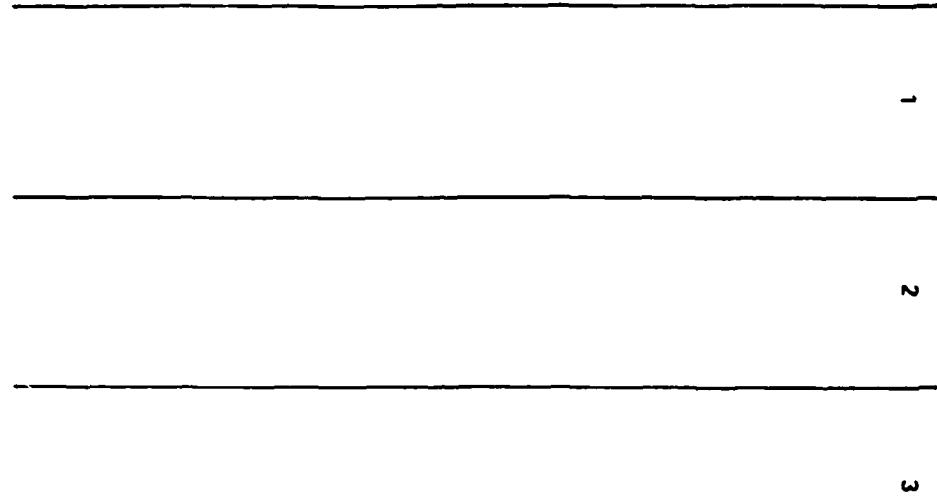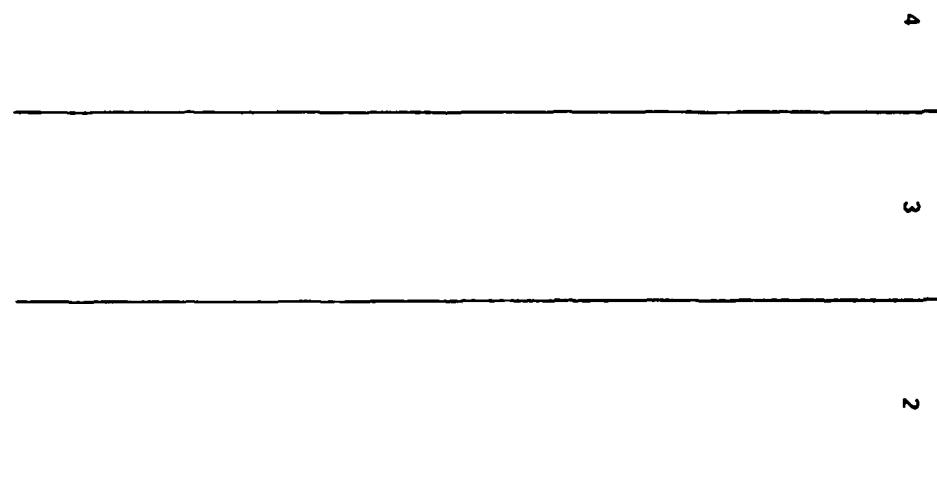
RIGHT CHANNEL BOUNDARY

Figure 6. Displacement

SECTION 3

RESULTS AND DISCUSSIONS

3.1 GENERAL

The principal objective of this screening experiment was to identify, from the six simulator characteristic variables investigated, those which have the greatest impact on the effectiveness of training. This would permit the planned subsequent experiments to fully investigate the training system/simulator design aspects of these most important characteristics, and hence reach definitive design conclusions. To accomplish this objective, the experimental training program had to be a constant for all experimental groups, except as modified by the experimental variables. This was accomplished through the careful design and implementation of the procedures discussed in Section 2.4. In addition, the training program had to be effective in order to establish a creditable basis for any decisions affecting the implementation of mariner simulator-based training. The training program did prove to be effective. Examples of its success are indicated by an overall pass-fail performance measure (i.e., pass: successfully transiting the scenario segment; fail: grounding or collision with traffic vessel). Only 8 percent of the trainees successfully completed leg 1 of the pretest while 74 percent successfully completed leg 1 on the posttest. Similarly, only 15 percent of the trainees successfully completed leg 4 on the pretest while 74 percent successfully completed it on the posttest (see Figures 7 and 8). Secondly, analysis of the debriefing interview data reveals a mean student rating of 9.23 out of 10 with regard to the effectiveness of the training program. This indicates an overwhelmingly favorable student reaction towards the training experiment, with a great majority of students requesting to be called to participate in future programs of this nature. This sentiment was also echoed in the students' evaluations of their own performances on the pretest and the posttest. It was found that there was a definite feeling of improvement due to the simulator-based training received. The mean self-rating on the pretest was 3.24, while the mean self-rating of the posttest was 7.84. Statistically, this difference was determined to be significant at the 0.001 level. The experimental training program was developed solely to support this experiment. It was

necessary to have an effective training program; the data indicates it was very effective. Moreover, it appears that this experimental training program had substantial practical value for advanced level deck officers as well.

3.2 LIMITATIONS OF SCREENING PROCESS

Several critical points should be kept in mind when interpreting the results of this screening experiment. First, the purpose of this screening experiment (particularly in the initial phases of the process) is to identify those variables (i.e., training system/simulator characteristics) that have the greatest relative impact on the effectiveness of training. That is, the purpose was to select the most important characteristics (e.g., perhaps three) from the six investigated; this would enable full investigation of the selected subset (e.g., the three most important characteristics). The least important characteristics would not be further investigated. This process is necessary to reduce the number of characteristics investigated to a size manageable within the resources available. The subsequent experiments would develop the detailed information from which the training system/simulator design decisions would be made. This report addresses the initial stage of the screening investigation to determine the relative importance of the six characteristics discussed in Section 2.2.2.

Second, the evaluation of experimental results and selection of the subset of important variables for more complete investigation, is an incomplete and imprecise process at this stage of screening process investigation. It requires the interpretation of several, and sometimes conflicting, analytical results to arrive at a best bottom-line estimate of relative importance regarding the six characteristics.

Third, in making the trivial/nontrivial decision with regard to a specific experimental factor, both the magnitude of the effect and the reliability of the effect should be considered. It should be noted that the tests of statistical significance only pertain to the reliability of an effect; the variance estimates the magnitude of the effect.

3-1
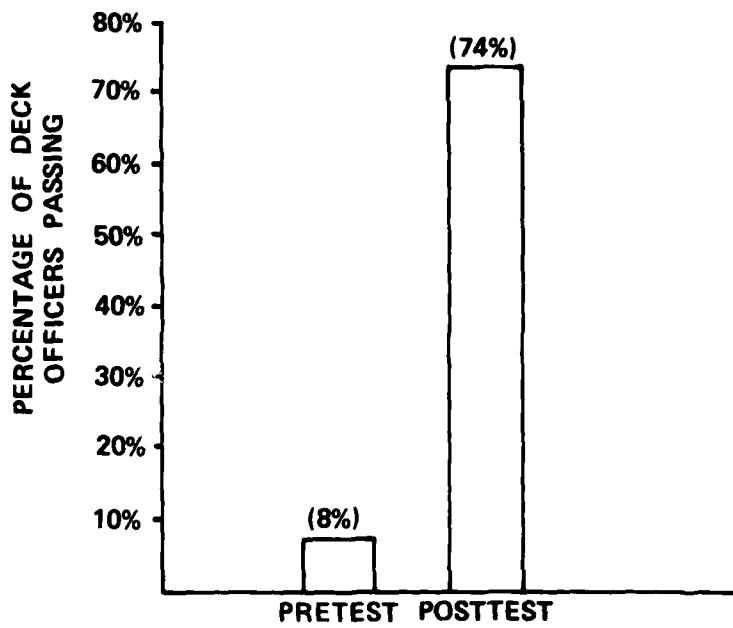
Figure 7. Percent of Successful Completion of Leg 1 for the Pretest and Posttest
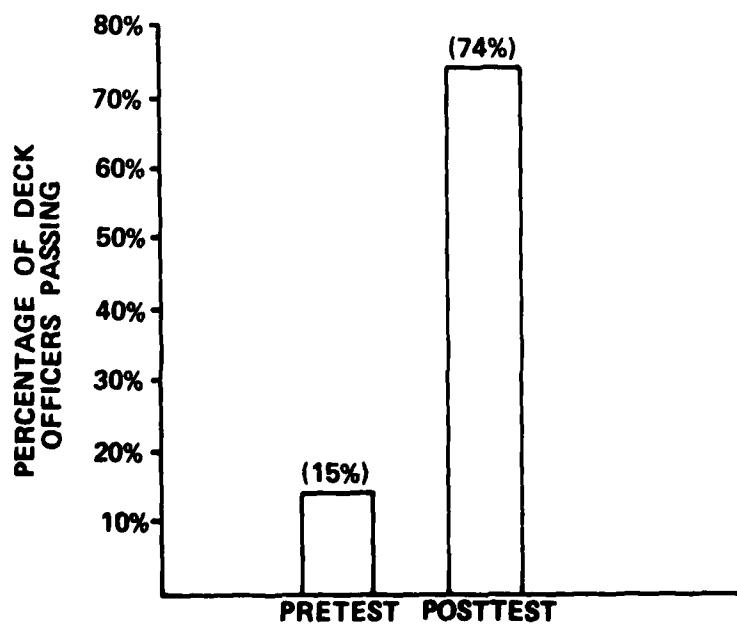
Figure 8. Percent of Successful Completion of Leg 4 for the Pretest and Posttest

Fourth, the results of the analysis are based solely on the levels selected for each experimental variable. The levels determine, to some extent, how each variable will appear to affect performance. If the levels are too close together there may be minimal performance variability. The investigator may read this (incorrectly) as an indication that the variable has a trivial effect on performance, when, in fact, this is true only within the levels selected or vice versa.

Fifth, the particular measure used to evaluate performance may influence the experimental results. Therefore, the analysis of these experimental data was conducted using multiple performance measures. (See Appendix J.) In addition, there appears to be no single optimum performance measure to analytically evaluate the performance of mariners in a restricted waters operational setting (test scenario).

3.3 FORMAT

The results of the analysis of the data from this screening experiment are presented and discussed individually for each of the experimental variables investigated. For each variable the following framework is utilized:

1. **Results.** A summary of the statistically significant results for both integrated and emergency shiphandling training objectives.

2. **Screening Process Interpretation.** A discussion of the importance of the variable with regard to simulator-based training and the rationale for retaining or disregarding it in future research investigations.

3. **Experimenters' Interpretation.** A discussion as to which level of the variable is preferable for simulator-based training and the authors' rationale for this preference. It should be noted that these "extended interpretations" are primarily based on subjective interpretation of the results. The objective of the screening process was to identify high importance variables that have significant impact on the effectiveness of simulator-based training. The identification of the proper level of the variable for simulator-based training was to be the subject of future investigations in the screening process. However, due to time and funding constraints which make the feasibility of future experiments questionable, the authors have been requested to project based on the available data the proper levels of simulator characteristics for each variable investigated.

3.4 TARGET MANEUVERABILITY

3.4.1 RESULTS

3.4.1.1 **Integrated Shiphandling.** There was only one statistically significant result with regard to the relevant performance measures investigated. In leg 1 the mean CPA of the groups trained with the canned targets was 59.94 feet greater than the mean CPA of the groups trained with the independent targets (F = 7.99; df = 1, 40; $p \leqslant 0.01$). In leg 4 the mean CPA of the groups trained with the independent targets was 30.07 feet greater than the mean CPA of the groups trained with the canned targets. However, this result was not statistically significant. Figure 9 indicates the magnitude of these posttest CPAs for both groups in both legs.

It should be noted that the CPA performance measure only utilized posttest scores since the majority of subjects in both test scenario legs did not successfully transverse the leg on the pretest to a point where a meaningful CPA with the traffic vessel could be obtained. As a result for CPA to indicate the effectiveness of training under canned and independent treatment conditions, the assumption must be made that the input characteristics, of the two groups, as regards the handling of traffic vessels, were equivalent.

3.4.1.2 **Emergency Shiphandling.** Target maneuverability was not investigated under emergency shiphandling conditions. There were no traffic vessels in the emergency shiphandling legs of the test scenario.

3.4.2 SCREENING PROCESS INTERPRETATION. Since a statistically significant result was observed under integrated shiphandling conditions using the relevant CPA performance measure, the target maneuverability should be considered for investigation in subsequent experiments of the screening process. It should be reiterated that in the screening process there are no mechanical means for selecting variables to be retained for future investigation. The researchers must make these decisions based on the available information. Since the identification of a variable as trivial removes it from future investigation, care should be exercised prior to discarding any variable. As discussed below, the interpretation of the results of this research appears to indicate that the level of target maneuverability does have an impact on the effectiveness of training integrated shiphandling skills. In addition, it should be noted that the target maneuverability variable under emergency shiphandling conditions still requires investigation. The final decision as to which variables should be investigated

LEG NO.1
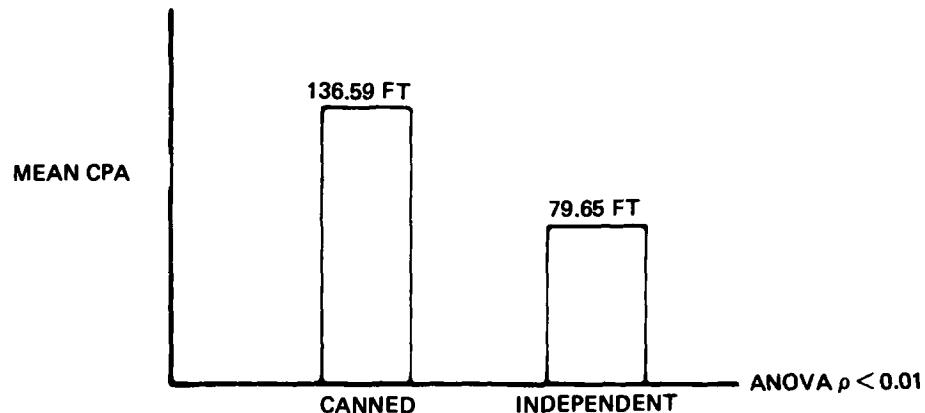
LEG NO.4
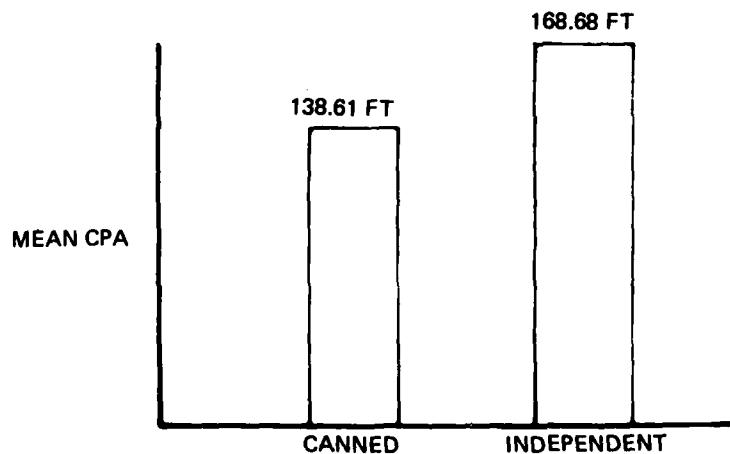
Figure 9. Mean CPA Comparison — Canned Versus Inde,endently Maneuverable Targets

in subsequent experiments depends on the relative findings between variables, and will be determined later in this report.

3.4.3 EXPERIMENTERS' INTERPRETATION. Target maneuverability was expected to be important in the handling of traffic vessels. The canned targets were expected to condition the subjects to anticipate certain traffic movements, resulting in significantly different CPAs compared to those subjects trained with independently maneuverable targets. The experimental results show that different CPAs did occur between the two groups in the two segments of the test scenario with traffic vessels (Figure 9). In leg 1, the canned group had a significantly greater CPA than the independent group ($F = 7.99$; $df = 1, 40$; $p \leq 0.01$). Conversely, in leg 4, the independent group had the greater CPA although the difference was not significant.

This dissimilarity in behavior between leg 1 and leg 4 may have resulted from the different degree of maneuvering difficulty and the different communications procedures used between the scenario legs. In leg 1 of the test scenario, the traffic vessel typically responded to ownship's call and

3-4

requested permission to pass ahead of ownship into a secondary channel (Figure 10). The independent group, having encountered similar situations in the training program, recognized it and responded accordingly. The subjects were familiar with the tug/barge occasionally cutting across their path to enter a north-bound channel. Hence, they were able to gauge an efficient/safe passing distance commensurate with their needs and past experience. The canned group, conversely, did not encounter the tug/barge ever cutting across their path; this request and subsequent action of the tug/barge, therefore, represented unusual behavior to which ownship was unaccustomed. Hence, in the canned situation, ownship appears to have been very cautious, achieving a greater CPA than that of the independent ownship. The larger CPA in this case is not necessarily better. Furthermore, if the mates in both groups are assumed to be equivalent and acceptable in shiphandling skill (i.e., both CPAs are acceptable), the independent situation group achieved a more cost effective performance.

The above rationale holds true for the behavior of both groups in leg 4 also, although conversely so for both groups. Whereas the test situation in leg 1 differed from the training situation for the canned group, the test situation in leg 4 was identical to the training situation experienced by the canned group (see Figure 10). Hence, the leg 4 situation was more familiar to the canned group than to the independent group. As a result, the canned group appears to have achieved a smaller mean CPA, although not significantly so; the independent group, on the other hand, was more cautious and achieved a larger mean CPA. That is, the canned group was most familiar with the movements of the traffic vessel in the test situation and thus achieved a more efficient, although still safe, CPA with the traffic vessel. Note, the rationale for leg 4 is based on mean CPA differences that were not significant; strictly speaking, both independent and canned groups did not differ in their leg 4 performance.

The greatest CPA, therefore, was apparently achieved by the group least familiar with the particular situation. The independent group encountered similar situations in legs 1 and 4, in that they had seen the traffic ship take different course of action in both legs. Hence, the independent group was likely to have exhibited the same degree of cautiousness in both leg 1 and leg 4. The canned group, on the other hand, encountered vastly different situations in legs 1 and 4. In the leg 1 test situation, the traffic ship behaved differently than what they had been accustomed to during the training program; in leg 4, however, the traffic

ship behaved exactly the same as they had been accustomed to in the training program. They were more likely to have exhibited a greater degree of cautiousness in leg 1 than in leg 4. The opposite direction of CPA differences support this interpretation.

The most effective configuration for training requires a further interpretation of these findings. If both efficiency of operation and safety are considered as factors in this determination, the independent group appears to provide a greater degree of training effectiveness. This interpretation is based on the fact that traffic ships will not always perform in a predictable manner, a necessary condition for more efficient behavior by the canned group. The canned group performance is likely to become less efficient under conditions when the traffic ships will not perform in the highly predictable manner. The independent group, on the other hand, is more accustomed to differences in behavior on the part of the traffic vessels. That is, they expect considerably more variation in behavior of the traffic vessels, and have learned to cope with it more efficiently. Hence, the efficiency of their behavior is less likely to vary as a function of normal differences in traffic ship actions. Assuming that safety is equivalent for the actions of both groups in legs 1 and 4, the independent group, therefore, would appear to have the more efficient behavior over all the likely situations.

In summary, this research appears to indicate that caution should be employed in the utilization of canned traffic vessels. There may be a danger that the use of canned traffic vessels during training may provide the mariner with a false sense of confidence in predicting the behavior of the other vessel. There are undoubtedly specific training objectives where the proper use of canned targets can provide the most cost effective training vehicle. However, until such guidelines are better defined, it would appear prudent to incorporate an independently maneuverable target capability into any simulator being considered for training skills involving interactions between two vessels, particularly in restricted waters.

3.5 COLOR VISUAL SCENE

3.5.1 RESULTS

3.5.1.1 Integrated Shiphandling. None of the performance measures utilized in the integrated shiphandling segments of the test scenarios (leg 1 and leg 4) revealed significant differences ($p < 0.10$) in training gain between those groups trained with the black and white visual scene and those

LEG NO. 1 TRAINING AND TESTING

TESTING

TRAINING
(CANNED)

TRAINING
(INDEPENDENT)
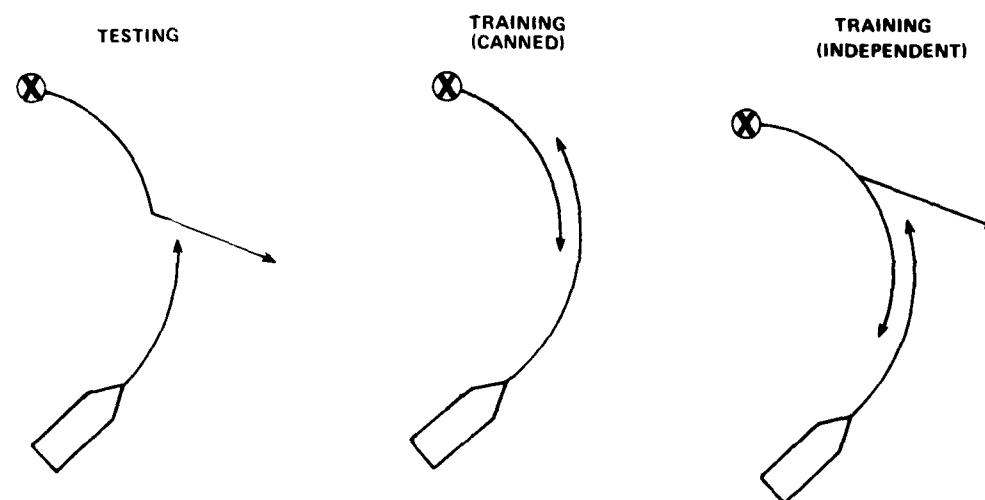
LEG NO. 4 TRAINING AND TESTING

TESTINC

TRAINING
(CANNED)

TRAINING
(INDEPENDENT)

KEY:

⊗ TRAFFIC VESSEL

◁ OWNSHIP
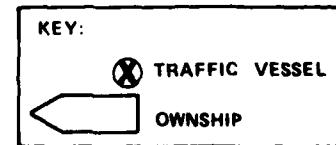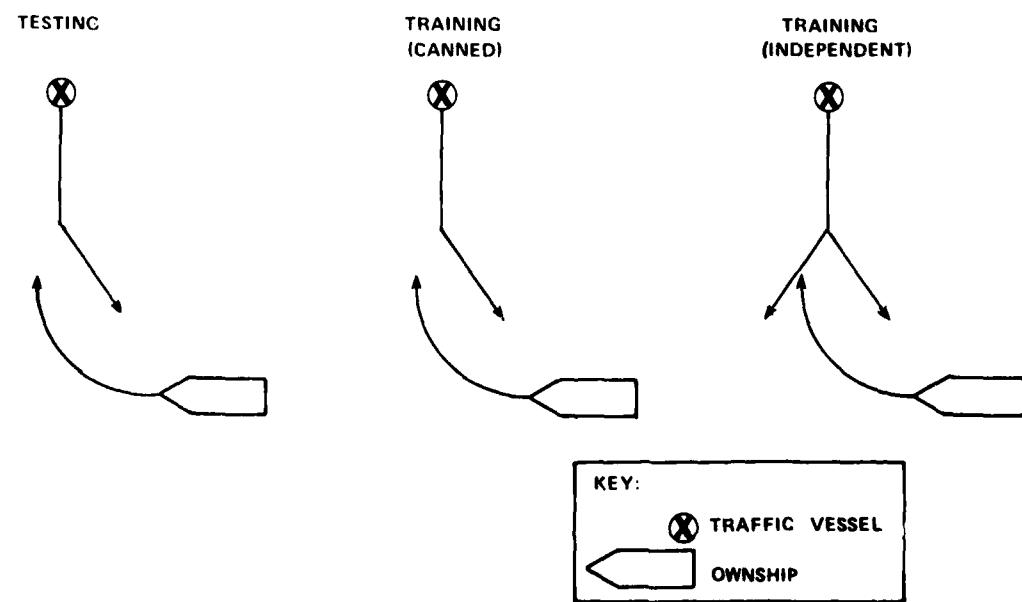
Figure 10. Training and Testing Sequences for Canned and Independently Maneuverable Targets

groups trained with the color visual scene. The low percentage of variance accounted for by the color visual scene variable and the low relative ranking within representative performance measures (Table 7) further indicates that the effect of this variable was trivial in the test scenarios employed.

3.5.1.2 Emergency Shiphandling. During the rudder failure segment of the test scenario (leg 2), those groups trained on the color scene had a mean improvement of 1.67 degrees in the vessel's heading towards the channel centerline. Those groups trained on the black and white scene had a mean degradation of 1.19 degrees in the vessel's heading away from the channel centerline. This difference was determined to be significant (F = 4.44; df = 1, 40; p ≤ 0.05). The deviation from desired heading was evaluated after the casualty as the vessel passed a preselected geographic station. During the main propulsion failure (leg 3), those groups trained on the color scene had a slight degradation in their mean deviation from the desired vessel heading of 0.48 degrees (i.e., measured as the absolute value from the channel centerline at preselected stations). Those groups trained on the black and white scene had a substantially greater degradation of 1.79 degrees. This difference was also determined to be significant (F = 3.4; df = 1, 40; p ≤ 0.07).

3.5.2 SCREENING INTERPRETATION. Although no significant results were observed for the integrated shiphandling legs and the practical meaning of the significant results observed in the emergency shiphandling legs (i.e., the difference between 0.48 and 1.79 degrees mean deviation from desired heading) may be argued by some readers, it is the researcher's judgment that the color visual scene variable should be retained for further investigation. The maritime community has a long history of utilizing different colors for vessel side lights and aids to navigation markings. To say that color is a trivial variable runs contrary to this tradition. It is more likely that for the shiphandling training examined in the particular test scenario was not sensitive to the color variable.

3.5.3 EXPERIMENTERS' INTERPRETATION. In daytime operations, the black and white presentation downgrades, but does not eliminate, any important visual cues. In nighttime operations, all lights became white, and the information transmitted by their color characteristics was provided via an associated flash code. This is not viewed as a problem with aids to navigation since it is possible to encounter, in the at-sea environment, geographic areas marked by only white lights (buoys and beacons) with distinctive flash rates. Hence, deck officers have experience in interpreting and using flash patterns. However, the side

TABLE 7. PERCENTAGE OF VARIANCE FOR EXPERIMENTAL VARIABLES —
INTEGRATED SHIPHANDLING
PM: Closest point of approach (CPA)
PM: Mean distance from recommended track (RT)

CPA				RT			
Leg 1	% Variance	Leg 4	% Variance	Leg 1	% Variance	Leg 4	% Variance
A	15.33	D	7.43	F	10.66	F	15.85
D	3.58	A	3.18	C	3.36	C	1.65
F	1.60	C	2.32	E	2.16	E	0.28
B	0.58	F	1.71	D	0.04	D	0.07
E	0.24	B	0.48	A	0.01	B	0.05
C	0.07	E	0.07	B	0.00	A	0.01

A   Target maneuverability   D   Time of day
B   Color visual scene   E   Field of view
C   Feedback methodology   F   Instructor

lights of traffic vessels must also be colored to be realistic. In this experiment, the red and green side lights of the traffic vessels were replaced with flash codes. This apparently did not hinder the performance of the groups trained under this condition.

The ability of the subjects to successfully train under conditions where they must process the flash rate of a light over time in lieu of instantaneously obtaining the red color of a port side light, may be due to the relatively light amount of traffic encountered in the scenarios. Their ability to keep track of traffic vessel movement was not taxed. Successful training may not have resulted if the same test subjects were placed in a port approach scenario with high contact workload where they would be required to more quickly identify traffic ship aspect. It should be noted that the test scenarios used to obtain the measure of each subject's training gain were daytime scenarios. It is assumed that similar results would be obtained if the test scenarios had been administered at night since the transformation from training with coded white lights to testing with colored lights should reduce (not increase) the temporal workload of the subjects. However, this assumption should be investigated by future research.

The negative training in leg 3 can be explained as the temporary degradation due to overreaction in the normal progression from the natural "do-nothing" reaction to the proper control actions (Johnson, 1976). In this leg, after the propulsion plant failure, the correct response is simply to steer the vessel until power is restored. As the vessel slows down, the rudder becomes increasingly ineffective. The proper mode for controlling the vessel's heading then becomes the bow thruster. The test subjects had little experience in the use of the bow thruster. On the pretest they neglected the bow thruster, while on the posttest they overused it. This theory is substantiated by the significantly different frequency of bow thruster usage (number of applications per minutes in run) between the pretest (0.1411) and the posttest (0.2085) ($F = 2.18$; $df = 4, 47$; $p \leqslant 0.01$). Due to the relatively short length of the training program, adequate training was not provided for the subjects to master this difficult skill.

The importance of the color visual scene variable in training emergency shiphandling skills is supported by the high relative ranking of this variable in percentage of variance explained. The ranking is based on several representative performance measures (Table 8 and 9).

TABLE 8. PERCENTAGE OF VARIANCE FOR EXPERIMENTAL VARIABLES — EMERGENCY SHIPHANDLING (RUDDER FAILURE)
PM: Distance from centerline (DL 40)
PM: Deviation from desired heading (DH)

Leg 2	Centerline	Leg 2	DH
E	6.0	E	18.30
B	4.5	B	7.80
C	3.0	A	0.88
F	1.1	F	0.10
D	0.7	D	0.02
A	0.1	C	0.00

A	Target maneuverability
B	Color visual scene
C	Feedback methodology
D	Time of day
E	Field of view
F	Instructor

TABLE 9. PERCENTAGE OF VARIANCE FOR EXPERIMENTAL VARIABLES — EMERGENCY SHIPHANDLING (POWER FAILURE)
PM: Mean deviation from desired heading (DH)
PM: Distance from centerline (DL 58)

Leg 3	DH	Leg 3	Centerline
B	7.26	C	7.30
D	4.10	B	1.70
F	1.30	A	0.27
E	0.35	F	0.13
C	0.19	E	0.03
A	0.01	D	0.00

A	Target maneuverability
B	Color visual scene
C	Feedback methodology
D	Time of day
E	Field of view
F	Instructor

It is not readily apparent why a color visual scene should be preferable for training in emergency shiphandling. Several plausibe explanations exist. First, the color visual scene provides an additional dimension of visual information beyond that of the black and white scene with flash coding. This additional dimension would enable the deck officer to process available information at a higher rate. The rate at which the deck officer would want to process information under normal conditions may be lower than that under emergency conditions. Hence, the black and white scene may present adequate information when the deck officer is not heavily loaded (i.e., under normal conditions); whereas, the color scene may afford him the higher information processing rate he desires under heavy loading (i.e., emergency conditions). In essence, the deck officer may be using, as rapidly as possible, all the relevant information he can acquire under emergency conditions.

A second explanation is that as a variable in the experiment, the color visual scene is aliased with two two-factor interactions: feedback/field of view interaction (CXE) and time of day/instructor interaction (DXF). One of these two interactions, and not the color visual scene variable alone, may have caused the effect. The specific cause must be determined by further research of the relevant variables.

In summary, this research appears to indicate that a color visual scene may not be required for some training objectives. However, until more specific guidelines are available to define the specific training objectives for which a black and white visual scene suffices, it appears only prudent to recommend a visual scene capable of simulating color for at least vessel side lights and aids to navigation — these being the principal color cues historically used by the maritime community.

3.6 FEEDBACK METHODOLOGY

3.6.1 RESULTS

3.6.1.1 Integrated Shiphandling. Inspection of seven relevant performance measures does not contain any significant differences (i.e., $p \leq 0.10$) between those groups trained with augmented feedback and those groups trained with nonaugmented feedback. However, in leg 1, five of the seven performance measures indicate greater training gain for the groups trained with the augmented feedback (see Table 10).

In leg 4, six out of the seven performance measures indicate a greater training gain for the groups trained with the nonaugmented feedback (see Table 11).

TABLE 10. TABULATION OF GREATER TRAINING GAIN FOR FEEDBACK METHODOLOGY LEVELS (LEG 1)

Performance Measure	Nonaugmented	Augmented
Deviation from centerline		●
Deviation from recommended track (RT)		●
Closest point of approach (CPA)	●	
Composite (centerline and CPA)		●
Composite (RT and CPA)		●
Composite (graphic)	●	
Pass-fail		

TABLE 11. TABULATION OF GREATER TRAINING GAIN FOR FEEDBACK METHODOLOGY LEVELS (LEG 4)

Performance Measure	Nonaugmented	Augmented
Deviation from centerline	●	
Deviation from recommended track (RT)	●	
Closest point of approach (CPA)	●	
Composite (centerline and CPA)	●	
Composite (RT and CPA)	●	
Composite (graphic)	●	
Pass-fail		

3.6.1.2 Emergency Shiphandling. The augmented feedback groups demonstrated significantly less deviation from the desired heading (i.e., 1.18 degrees more, $p \leq 0.1$) in leg 2 (rudder failure). The augmented feedback groups also demonstrated less degradation in distance from channel centerline (i.e., at data line 58; 54.68 feet less degradation, $p \leq 0.1$).

3-9

3.6.2 SCREENING INTERPRETATION. As a result of the apparently conflicting trends identified for the integrated shiphandling training along with the statistically significant results identified for the emergency shiphandling training, the feedback methodology variable should be retained for investigation in future experiments of the screening process. Previous research has shown feedback displays to be an effective supplement to the simulator-based training process (Hammell, et al., 1978). The U.S. Navy is currently conducting an extensive development and evaluation of many aspects or advanced training technology (Hammell, et al., 1980a; N61339-80-C-0079). Preliminary indications are that this technology should greatly impact training effectiveness. The absence of a strong effect for the feedback variable in this experiment may be the result of several factors including the improper matching of the feedback device with the training objectives as discussed in the Section 3.6.3.

3.6.3 EXPERIMENTERS' INTERPRETATIONS. The presence of the CRT display showing a "real time" plan view of ownship's transit of the channel was of assistance to trainees in the first segment of the waterway, and appears to have been an impediment to performance in the final leg. The explanation of this difference, the reversal of greater training effectiveness, is based on a comparison of the characteristics of the feedback display, with the nature of the shiphandling problem in the two integrated shiphandling segments. The feedback display consisted of a plan view of the waterway showing ownship's position with respect to channel boundaries and traffic vessels. Ownship's position with respect to traffic vessels was not represented to scale on the CRT screen, but was represented as point sources with speed/heading vectors.

In leg 1, the "integrated" nature of the shiphandling problem consisted of making a difficult channel bend under conditions of considerable current shear and set, followed by a meeting of a single outbound traffic vessel. As indicated by (1) the pretest frequency of right channel boundary excursions immediately following ownship's experience with the current and, (2) the fact that very few trainees made it through the bend and up to the traffic ship in pretest runs indicates that this was without doubt a difficult scenario which needed to be addressed in training.

The feedback display would appear to be an effective aid in assisting the students to acquire the necessary skills to properly negotiate leg 1. It afforded a plan view of the scenario which, in contrast to the perspective view through the bridge windows, provides a more immediate, graphic

and precise appreciation of ownship's cross channel displacement. Additionally, the feedback display depicts channel boundaries as solid lines giving continuous "distance off" information to trainees, as contrasted with the visual scene through the bridge windows which shows only intermittent buoys marking an estuary approach channel that is not bounded by shoreline contours. Therefore, the addition of the feedback display should be of positive value in training for the application it gave regarding the considerable effect of the current and the need for early and ample control action.

One performance measure which did not indicate greater training effectiveness due to employment of feedback display was that of CPA. The reasons for this lack of feedback display training effectiveness in traffic ship related performance measures are more apparent in an examination of leg 4. Leg 4, also an integrated shiphandling scenario, differed from the first leg of the exercise in that: (1) cross channel displacement due to current effect was not so pronounced, (2) the channel limits were clearly paralleled by distinctive shoreline banks, and (3) three vessels rather than a single traffic ship was encountered. Consequently, for this fourth leg, feedback display cross channel positioning information was not appreciably better than that afforded by the out-the-window view, where continuous "distance off" information was available from comparative observation of the opposite shore lines. Also, cross channel position information was not as essential as it was for the more difficult current conditions of leg 1. Current posed less of a problem in the turn of leg 4 possibly because of the characteristics of the flow being absolutely less difficult to contend with, or perhaps because of the subjects' pretest "training" in negotiating the bend of leg 1.

Regarding the handling of traffic ships, the nonaugmented groups appear to have achieved greater CPAs than the augmented groups. One possible explanation of this result may be that the immediate feedback display used during the experiment contained insufficient information on the traffic ship. The traffic ship's position and velocity vector supplied by the feedback display were substantially less informative for the restricted water scenarios employed than the traffic ship information presented by the visual scene. The traffic ship element of the integrated shiphandling problem was much more pronounced in leg 4 than in leg 1. Two stationary tugs served as obstructions in the outer bend of the leg 4, turn, whereas none were present in leg 1. Secondly, the moving traffic ship of leg 4 did not come into view as early as it did in leg 1, and passed

ownship in the middle of the turn rather than afterward. Use of the visual scene allows a better and a continuous evaluation of the changes in orientation (heading) of outbound traffic ships than did the plan view of the feedback display which is much less able to convey this type of information. Evaluation of moving traffic ship problems requires "target angle" rather than "distance off" information. Of the two situation displays or "views," the perspective rather than plan presentation is therefore most appropriate. Apparently, the presence of the feedback display during training, which contributed very little due to (1) its limitations and (2) the nature of the shiphandling problem in the final turn, competed with the bridge window visual display for the attention of the subjects. Time spent watching the CRT display equaled time not invested in evaluating the angular approach of the moving traffic ship. The particular feedback display chosen, therefore, was apparently inappropriate for the leg 4 situation; it was, however, of apparent value in training regarding the leg 1 situation. A summary discussion of training technology and its application on training devices is presented in Appendix M.

The results of this variable evaluation indicate that feedback displays must be designed and employed with regard to specific training objectives. A single display will not meet all training needs, especially in integrated shiphandling exercises in which scenarios address the development of quite different perceptual and vessel control skills. The augmented feedback may be more effective, but only when tailored to the specific training objectives.

3.7 TIME OF DAY

3.7.1 RESULTS

3.7.1.1 Integrated Shiphandling. In leg 1 based on the pass/fail performance measure, 18 of the subjects trained under day improved; whereas 13 of the subjects trained under night improved (Fisher Test; $p \leqslant 0.03$). In addition, four of five performance measures which incorporate CPA information indicated that the subjects trained under day had greater training gains (see Table 12). This may indicate that ambient lighting conditions have an impact on the effectiveness of training skills which involve the interaction of two vessels in restricted waters (see Section 3.7.2). In leg 4, 16 of the subjects trained under night improved, whereas 12 of the subjects trained under day improved (Fisher Test; $p \leqslant 0.02$). Also in leg 4, the mean CPA of the groups trained at day was 45.95 feet greater than the mean CPA of the groups trained at night (ANOVA; $p \leqslant 0.08$).

TABLE 12. TABULATION OF GREATER TRAINING GAIN FOR TIME OF DAY LEVELS (LEG 1)

Performance Measure	Night	Day
Closest point of approach (CPA)		●
Composite (centerline and CPA)	●	
Composite (RT and CPA)		●
Composite (graphic)		●
Pass-fail		●

Caution should be utilized when interpreting the results of the pass-fail performance measure in leg 4 for this variable since this performance measure inherently assumes equivalent input proficiency between groups which did not exist in leg 4. Only one subject in the night group passed the pretest while six subjects in the day group passed. This imbalance means that for analysis of this variable using the pass-fail performance measure in leg 4 the day group had considerably less potential for improvement than the night group.

3.7.1.2 Emergency Shiphandling. The only significant result for emergency shiphandling training occured in leg 2 during the rudder failure utilizing the pass/fail performance measure. Six of the subjects trained under day improved whereas 2 of the subjects trained under night improved (Fisher; $p \leqslant 0.08$).

3.7.2 SCREENING INTERPRETATION. Based on the results cited in the previous section, it appears that the time of day variable may have a significant impact on the effectiveness of simulator-based training. However, the results of this research and the interpretation discussed below do not provide comprehensive or conclusive answers for the long standing debate concerning the necessity for a day/night simulator over a night only simulator (although additional insight is gained). The time of day variable should be retained for future investigations in the screening process.

3.7.3 EXPERIMENTERS' INTERPRETATION. In analyzing the results of this variable (time of day), it should be noted that although some groups were trained under day conditions and other groups were trained under night conditions, all groups were tested under day conditions. The results of this experiment, therefore, appear to indicate that daytime training is preferable for daytime operations.

This finding contradicts the hypothesis that only nighttime simulator based training should be required since nighttime operations are more difficult than daytime operations, and therefore, training for nighttime operations ensures adequate training for daytime operations. Nighttime operations may be more difficult than daytime operations. However, different skills, or different application of the same skills, may be required for daytime operations. For example, the skill required for perceiving the traffic ship's aspect in a meeting situation at night may be an analytical skill based on the direction and distance between the forward and after masthead lights; while for daytime operations the skill may be a perspective skill using the relative difference between the portside area of the traffic vessel and the starboard side area. Hence, the training requirements appear to differ between night and day, suggesting separate training.

Based on the results of this experiment and the rationale that nighttime operations require different skills than daytime operations, it is logical to assume that nighttime training would be preferable for nighttime operations. However, further investigations should be conducted to evaluate the effect of daytime versus nighttime training for nighttime operations.

One plausible interpretation as to the reason that daytime training was better for responding to the rudder failure may have been related to the perceptual skill that these individuals developed for sensing the movement of the vessel towards the bank after being notified of the rudder failure. Although the response to this casualty was procedural, the shiphandler must confirm that the rudder is jammed, not just the rudder angle indicator, prior to executing the response. Those individuals trained during daytime may have developed better skills for sensing motion towards the bank during daytime operations than those trained at night. Once again, although daytime training appears preferable, it may be unsatisfactory for night operations.

Finally, it should be noted that although nighttime training may not be as effective as daytime training for daytime operations, it still may provide acceptable training for daytime operations. However, based on the information available today, it appears prudent to train shiphandling skills under the ambient lighting conditions that will be utilized at sea. As a result, this appears to imply a day/night capability for training facilities offering a comprehensive simulator-based training program.

3.8 HORIZONTAL FIELD OF VIEW

3.8.1 RESULTS

3.8.1.1 Integrated Shiphandling. Inspection of seven relevant performance measures for leg 1 reveals one acceptably significant finding. From pretest to posttest, the groups trained with 120 degree field of view were 59.61 feet closer towards the channel centerline than the groups trained with 240 degree field of view ($F = 3.32$; $df = 1, 40$; $p < 0.08$). Figure 11 illustrates the magnitude of the training gains for both of these groups. There were no significant findings in leg 4 for this variable.

Inspection of the relevant performance measures also reveals that for five out of the seven performance measures, the groups trained with 120 degree field of view had greater training gains in leg 1 than the groups trained with 240 degree field of view (Table 13). In leg 4, five out of the seven performance measures indicated a greater training gain for the groups trained under the 240 degree field of view (Table 14).

3.8.1.2 Emergency Shiphandling. Inspection of six relevant performance measures for leg 2 revealed four significant findings which are summarized in Table 15. From pretest to posttest, the groups trained with 120-degree field of view improved their desired heading at data line 40 by 4.38 degrees more than the groups trained with 240-degree field of view ($F = 10.39$; $df = 1, 40$; $p < 0.003$). From pretest to posttest, with regard to the pass-fail performance measure, seven subjects trained with the 120-degree field of view improved whereas only one of the subjects trained with the 240-degree field of view improved (Fisher; $p < 0.02$). For the composite (graphic) performance measure, the score for the students trained with the 120-degree field of view (574.0) was significantly greater than the score for the students trained with the 240-degree field of view (461.0) (Mann-Whitney V; $p < 0.06$). From pretest to posttest, the groups trained with the 120-degree field of view improved 1.12 degrees more than the groups trained with the 240-degree field of view ($F = 2.95$; $df = 1, 40$; $p < 0.09$).

Inspection of five relevant performance measures in leg 3 reveals no significant findings. However, two other nonsignificant observations should be noted. First, the following performance measures indicate slight negative training gains (i.e., training losses): (1) deviation from channel centerline, (2) deviation from desired heading, and (3) deviation from centerline at data line 58. Second,
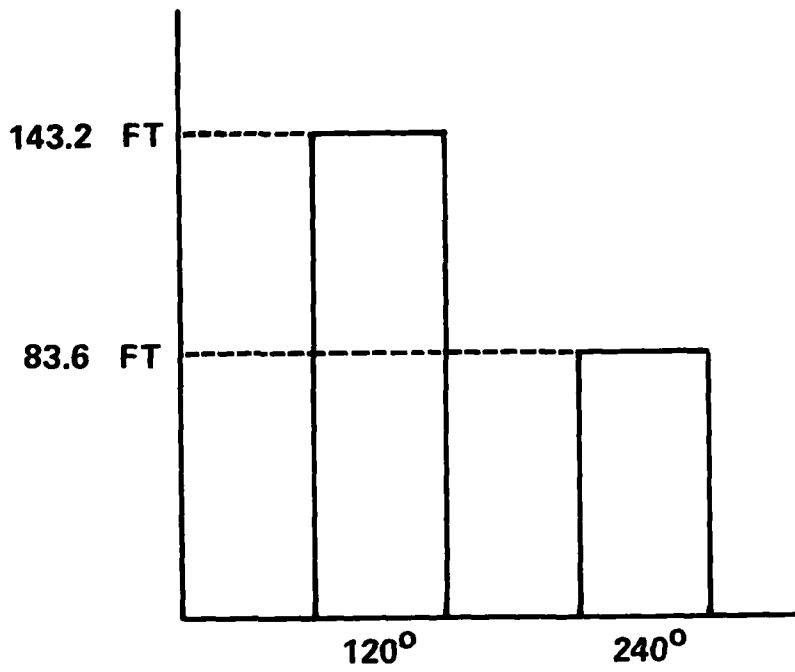
Figure 11. Magnitude of Training Gain for Horizontal Field of View Levels Investigated

TABLE 13. TABULATION OF GREATER TRAINING GAIN FOR HORIZONTAL FIELD OF VIEW LEVELS INVESTIGATED (LEG 1)

Performance Measure	120 Degrees	240 Degrees
Deviation from centerline	●	
Deviation from recommended track (RT)	●	
Closest point of approach (CPA)	●	●
Composite (centerline and CPA)	●	
Composite (RT and CPA)	●	
Composite (graphic)	●	
Pass-fail	●	

TABLE 14. TABULATION OF GREATER TRAINING GAIN FOR HORIZONTAL FIELD OF VIEW LEVELS INVESTIGATED (LEG 4)

Performance Measure	120 Degrees	240 Degrees
Deviation from centerline		●
Deviation from recommended track (RT)		●
Closest point of approach (CPA)		●
Composite (centerline and CPA)		●
Composite (RT and CPA)		●
Composite (graphic)	●	
Pass-fail	●	

3-13

TABLE 15. SUMMARY PERFORMANCE MEASURE DATA SHEET: FIELD OF VIEW (LEG 2)

Performance Measure	Variable: Field of View		
	120 Degrees	240 Degrees	Significance
Deviation from desired heading	1.17	0.05	$p \leqslant 0.09$
Deviation from desired heading (DL 40)	2.43	-1.94	$p \leqslant 0.003$
Composite (graphic)	574.0	461.0	*$p \leqslant 0.06$
Pass-fail	7/15	1/23	**$p \leqslant 0.02$

*Mann-Whitney U Test
**Fisher Test

four out of the five performance measures indicate greater training gain (or less training loss) for the groups trained with 240-degree field of view. See Table 16.

3.8.2 SCREENING PROCESS INTERPRETATION. Findings for a horizontal field of view indicate that each of the two fields investigated (i.e., 120 and 240 degrees) resulted in more effective training under certain circumstances. That is, neither field of view resulted in a clear-cut greater impact on the effectiveness of training; rather, each was more effective for certain segments of the training and test situation. Furthermore, these apparent interactive effects appear to be somewhat independent of the integrated shiphandling or emergency shiphandling situations. Both of these situation categories showed preference for each field of view under different circumstances. It appears, therefore, that the horizontal field of view is a variable that should receive careful consideration in the design of a ship bridge simulator. It is important to note that these results

TABLE 16. TABULATION OF GREATER TRAINING GAIN FOR HORIZONTAL FIELD OF VIEW LEVELS INVESTIGATED (LEG 3)

Performance Measure	120 Degrees	240 Degrees
Deviation from centerline		●
Deviation from desired heading		●
Deviation from centerline (DL 58)		●
Composite (graphic)		●
Pass-fail	●	

indicate a preference for the lower fidelity field of view under certain circumstances. These findings are obviously of substantial importance since (1) they may have a substantial impact on the training device cost, and (2) they indicate an actual training effectiveness preference for the lower cost/lower fidelity level under certain circumstances. Further investigation during subsequent screening process experiments is definitely warranted for this field of view variable.

3.8.3 EXPERIMENTERS' INTERPRETATION. The importance of horizontal field of view in mariner simulator-based training has been widely discussed. In this experiment the integrated shiphandling test scenarios (leg 1 and leg 4) were designed to be sensitive to this variable. The simulated vessel makes a 59-degree turn in leg 1 and a 129-degree turn in leg 4. It should be noted that the 120-degree field of view equates to a visual scene spanning ±60 degrees on either side of the vessel centerline. The criticality of this variable in these scenarios is understandable. However, the indication that more effective training occurs under the 120-degree field of view in leg 1 is initially startling since the more effective training would normally be thought to be equated to the higher fidelity (240-degree field of view) condition. One explanation of this trend may be that the successful trainee in the leg 1 scenario develops his perceptive skills with regard to visual cues in the geographic area directly ahead of the vessel. The reduced field of view prompts the trainees to concentrate their attention on this geographic area, hence, more effectively developing their perceptive skills with regard to the more critical visual cues. The turn magnitude (129 degrees) in leg 4, on the other hand, may cause the trainees to require information beyond the 120-degree presentation, hence, the indication of greater training effectiveness for the 240-degree field of view.

The rationale presented above for field of view differences may also apply to the more effective training observed in leg 2 under the 120-degree field of view. The emergency condition in leg 2 was that of a rudder failure during which the deck officer would likely be most concerned with the heading of ownship. Again, the more narrow field of view would encourage the deck officer to focus his attention ahead of ownship as opposed to dealing with information abeam. The significant findings on three of the four legs in the test, therefore, can support this interpretation. Several of the magnitudes of improvement on individual measures may not be particularly of consequence; nevertheless, the consistency of results obtained across several measures does substantiate the potential effectiveness of the lower field of view. The effectiveness of field of view depends upon specific aspects of the training situation; hence, both 240 and 120 degrees were more effective under certain circumstances.

The greater training gain in selected scenarios for the groups trained with the reduced horizontal field of view may be beneficial for use in training specific skills. It adds creditability to the part-task training approach in which the student is trained under limited scope simulations (i.e., reduced fidelity simulation) to promote the development of only certain specific skills. The student then progresses from the relatively inexpensive part-task training device to the full-mission training device where he can integrate the different skills which he has acquired separately. Such an approach to training, if utilized properly, has the potential to provide more cost effective training as it allows the more expensive full-mission simulator to be used primarily for those aspects of training requiring more complete simulation, such as task integration. This allows the training facility to increase the number of students that it can train on its full-mission simulator during a given training period (Miller, 1974; Lumsdaine, 1960).

Another implication of this result (i.e., greater training gain with reduced field of view in selected scenarios) may be to incorporate an adjustable field of view on the full-mission simulator. Such a horizontal field of view could be tailored, as appropriate, to more effectively train selected skills during the training process. This approach would be utilized principally for increasing the training gain, the rate at which the student would be trained to the standard level of performance. This would decrease the time required to train and hence increase the cost effectiveness of training for certain objectives. Care, of course, must be exercised to ensure the simulator characteristics would be manipulated

most effectively to achieve both the intermediate and long-term training objectives.

This finding (i.e., the lower fidelity is more effective in certain situations), although unexpected, is logical. The reduced field of view caused the trainees to focus attention on the important situation information, thus achieving a greater training gain for that particular situation. The greater training effectiveness of lower fidelity characteristics is a pioneer finding, although researchers have alluded to this possibility. It represents a classic difference between experience which is encumbered by all the real-world distracting factors (e.g., irrelevant information), and training, which can focus on the relevant factors in a structured progressive approach to achieve more effective and efficient skill acquisition. It should be cautioned that the use of restricted field of view, or other similar techniques, to enhance training should be done only in the context of a progressive training program that later addresses performance under conditions of distracting information. That is, for example, the restricted field of view may be used in the early stages of training to shape deck officer behavior to focus on the important information ahead of own ship. It is likely to be necessary, however, that later stages of training should address the filtering of the additional irrelevant information normally encountered at sea with a wider field of view. The later stages of training, therefore, should use the wider field of view. This type of a progressive training approach should achieve the overall training objectives more efficiently, than always training with the wider field of view. It builds the skills in an effective progressive fashion.

3.9 INSTRUCTOR

3.9.1 RESULTS. Variability in instruction was intended to be held constant through the use of a single instructor and a highly structured course, as this experiment was principally an analysis of the effects of simulator design characteristics on mariner training effectiveness.

However, the logistics of the experiment required the use of more than one instructor. Since it was believed that instructor differences have the potential to result in high variation of student performance, a second instructor was incorporated into the experiment through a balanced design, in order to maintain the accountability of the experimental effects. Hence the "instructor" became an independent variable at two levels, instructor "A" and instructor "B." The levels, however, were not intended to represent any particular characteristic of instructor qualifications, both individuals being chosen on the basis of acceptable shiphandling and teaching experience.

The results of the experiment have justified such caution in that the instructor variable was the most critical of all variables investigated.

3.9.1.1 Integrated Shiphandling. Numerous statistically significant results were observed for the instructor variable as a result of the integrated shiphandling training. The following results show greater average improvement under instructor "A" than under instructor "B."

- 80.95 feet closer toward the channel centerline in leg 1 from pretest to posttest (ANOVA; $p \leq 0.02$).

- 98.95 feet closer towards the recommended track from pretest to posttest in leg 1 (ANOVA; $p \leq 0.03$).

- A mean score of 62.56 points higher for the composite (mean distance from channel centerline and CPA) performance measure (ANOVA; $p \leq 0.06$).

- A mean of 80.57 points greater for the composite (mean distance from recommended track and CPA) performance measure (ANOVA; $p \leq 0.09$).

In addition to the above results, a relatively large percentage of variance was accountable to the instructor variable for many of the performance measures. For example, in Table 17 using the mean distance from recommended track performance measure, the instructor variable accounted for 10.66 percent of the experimental variance in leg 1; this is more than that of all other variables combined for that leg. Likewise, in leg 4 the instructor variable accounted for 15.85 percent of the experimental variance which is considerably more than that of all other variables combined. It may be noted that the experimental variables account for only a relatively small proportion of the total experimental variance. This is due to the nature of this screening design in which multiple variables are simultaneously manipulated, resulting in the presence of numerous interaction effects and adding to the complexity of the scenario situations. The identification of the variances associated with these interaction effects requires the conduct of additional experiments. The exploratory nature of the screening process recognizes and compensates for the infeasibility of identifying all sources of variance (see Appendix F).

3.9.1.2 Emergency Shiphandling. There were two statistically significant findings as a result of the emergency shiphandling training. In leg 2, based on the pass-fail performance measure, 6 of the subjects trained by instructor "B"

TABLE 17. PERCENTAGE OF VARIANCE FOR EXPERIMENTAL VARIABLES — INTEGRATED SHIPHANDLING

PM: Mean distance from recommended track (RT)

Leg 1	RT	Leg 4	RT
F	10.66	F	15.85
C	3.36	C	1.65
E	2.16	E	0.28
D	0.04	D	0.07
A	0.01	B	0.05
B	0.00	A	0.01

A	Target maneuverability
B	Color visual scene
C	Feedback methodology
D	Time of day
E	Field of view
F	Instructor

improved whereas 2 of the subjects trained by instructor "A" improved (Fisher; $p \leq 0.10$). In leg 3, from pretest to posttest, the groups trained by instructor "B" were 54.68 feet farther away from the channel centerline than the groups trained by instructor "A" (ANOVA; $p \leq 0.06$). Degradation of performance in leg 3 is discussed in Section 3.5.

3.9.2 SCREENING INTERPRETATION. The results of this screening experiment indicate that the instructor is an extremely important variable in the development of an effective simulator-based training program. The importance of a well qualified instructor and other intangible training program aspects has been well known for many years (Caro, 1973; Charles, 1976). The instructor was, by far, the most important variable investigated in this experiment. That is, the instructor had the greatest impact on the effectiveness of the training program. The results show that both instructor "A" and instructor "B" excelled in certain respects. Since no attempt was made to control the characteristics of the instructors for experimental purposes, conclusions regarding instructor characteristics are inappropriate at this time. Nevertheless, the results indicate the importance of the instructor to the simulator-based training system. This variable should be thoroughly investigated in subsequent experiments of the screening process.

3.9.3 EXPERIMENTERS' INTERPRETATION. The apparently greater training gain for the groups trained by instructor "A" in integrated shiphandling does not necessarily indicate that instructor "A" was the better instructor. It should be noted that the experimental results verify that both instructors were effective. However, examination of the correlation coefficients between pretest and posttest scores for the experimental groups reveals a significant number of negative correlations. A negative correlation coefficient indicates that those students who scored lowest on the pretest scored highest on the posttest. This is independent of training gain or loss. Further analysis revealed that instructor "B" training groups showed a higher percentage of negative correlation coefficients than those of instructor "A" (Table 18). This could indicate a difference in instructional strategy. Instructor "B" may have focused his attention on the training needs of the students entering with the least skills (as evidenced by pretest performance) while comparatively neglecting the development of "finesse" in those students already possessing a higher level of basic shiphandling skills when entering the course. As a result, instructor "B" was probably more effective in raising the students in the lower portion of his class (as measured on the pretest) above a minimum proficiency level than instructor "A." This may be an important concept if the principal thrust of mariner simulator-based training is to improve the minimum acceptable proficiency levels for licensed deck officers. In summary, therefore, these findings show that the instructors differed, but they do not provide conducive information regarding the desirability of specific instructor characteristics for deck officer training.

It may be noted that the effect of the instructor variable does not appear as pronounced in the emergency shiphandling legs as it was in the integrated shiphandling legs. This does not necessarily mean that the instructor variable is less important in emergency shiphandling. No effort was made to select the two instructors to represent two distinct levels of the instructor variable. Important characteristics that reflect the instructor's ability to train the proper procedural response associated with emergency shiphandling may not have differed between instructors. In order to gain better understanding of the instructor variable with regard to emergency shiphandling skills, additional investigations are required.

The instructor was included as a variable in this experiment out of necessity rather than as an objective to investigate the impact of the instructor on training effectiveness. The training literature generally supports the notion that the instructor may have a substantial impact on the effectiveness of any training program, despite the aids provided for his use. The simulator, of course, is merely an aid to the instructor conducting the training program. Hence, the overwhelming impact of the instructor, more so than that of any other characteristics investigated, is not unexpected. Relatively little, however, can be said about those characteristics that result in an effective instructor since this experiment was not designed to investigate them. Some information is available from the research literature; a brief summary of instructor attributes is presented in Appendix L. A summary of instructor characteristics for deck officer training is presented below.

In the research literature, three primary categories of characteristics have been associated with effective instructors. These are (1) expertise in the subject matter area being addressed (e.g., shiphandling expertise), (2) expertise in teaching (i.e., instructional expertise), and (3) acceptability to the trainees (e.g., a licensed deck officer).

TABLE 18. PRETEST/POSTTEST SPEARMAN CORRELATION COEFFICIENTS

PM: Mean Distance from Channel Centerline

Group	Instructor "A"	Group	Instructor "B"
Leg 1			
8	−0.02	6	−0.51
5	−0.41	4	−0.41
7	+0.57	1	−0.41
2	+0.85	3	−0.16
	$\bar{r} = 0.24$		$\bar{r} = 0.37$
	$\sigma = 0.56$		$\sigma = 0.15$
Leg 2			
8	−0.03	6	−0.09
5	+0.74	4	−0.51
7	+0.68	1	−0.37
2	−0.09	3	−0.66
	$\bar{r} = 0.33$		$\bar{r} = -0.41$
	$\sigma = 0.44$		$\sigma = 0.24$

An instructor having appropriate qualifications in each of these categories is likely to be an effective instructor. The first and second categories are relatively straightforward. That is, the instructor should have expertise in the subject matter he is teaching, for example shiphandling, to the level required for the particular course. That level will usually be substantially beyond the skill level standards expected as output characteristics of the trainees successfully completing the course. The necessary level of expertise by the instructor is dependent on a variety of factors including depth of expertise in a particular area as well as breadth of expertise. The expertise is required for demonstration to the trainees, for adequate explanation to the trainees, and for the development of appropriate analogies (i.e., examples) to assist their understanding. The requisite level of expertise will ensure that the instructor can, in fact, impart the necessary information to the trainees in a meaningful and effective manner. It should be noted that shiphandling expertise does not necessarily require a license. Furthermore, the level of expertise would definitely depend on the level of the trainees.

The level of expertise as an instructor should be viewed in a manner similar to that for subject matter (e.g., shiphandling). That is, not only must the effective instructor know the material and possess the necessary shiphandling skills to conduct an effective training process, but he must also know how to train effectively. He must understand and apply effective training methodology in conducting the training process. This is particularly important if he is provided with a variety of training aids, such as a simulator. These are expected to increase the cost effectiveness of training; they may, however, decrease the cost effectiveness if not used appropriately. As was noted above regarding subject matter expertise, the level of instructional expertise can vary with the level of the students as well as other aspects of the training system and the training situation. Nevertheless, these must be carefully determined in selecting and preparing the instructor. It should be noted that the area of instructional expertise is generally overlooked when selecting an instructor. It is expected that an individual having operational expertise will also be an effective instructor. An effective instructor would be one who has had considerable experience and/or instruction in the methodologies and skills of conducting an effective training process. Hence, this category of characteristics must receive careful consideration in the selection of an instructor and in the development of training system acceptance criteria.

The third category, acceptability to trainees, is necessary to establish a level of rapport with the students. The acceptability of an instructor does not depend upon any particular qualifications although holding a deck officer's license of an appropriate level would certainly be an advantage in gaining the acceptance of the trainees. Generally speaking, instructors holding a deck officer's license would more likely be acceptable to the trainees than nonlicensed instructors. Likewise, an instructor holding a master's license would be more acceptable to master's level trainees than an instructor holding only a 3rd mate's license. Nevertheless, holding a license and the level of that license would not ensure acceptability of the instructor by the trainees. Nor would the absence of a deck officer's license necessarily result in unacceptability of the instructor by the trainees. This is a complex category and one that is often dependent upon the personality aspects of the individual instructor along with his background experience, etc.

The above three categories of characteristics pertain directly to the instructor. The effectiveness of the instructor is heavily dependent upon the specificity of the training objectives he is to achieve, the materials he is provided which assist him in directing the training process and which provide him the necessary information for its conduct, and the tools at his disposal to provide feedback information, illustrate relationships, etc. These nonsimulator training program characteristics directly impact the effectiveness of the instructor conducting the training process. They too must be adequately specified and included in the training system design to augment the instructor in achieving an effective training process. The discussion regarding the feedback methodology characteristic investigated in this experiment directly addresses training assistance technology in support of the instructor. In addition to this technology, the specificity of the training objectives, the adequacy in detail level of the training curriculum material developed (e.g., instructor's guide, student handout material), classroom visual aids (e.g., slides), and the direction given to the training program (e.g., bring all deck officers up to a minimal skill level versus increase the skill level of all deck officers) are elements of the nonsimulator-related aspects of the training system that are closely integrated with the instructor's effectiveness.

The U.S. Coast Guard should give careful consideration to several issues in the development of the training system acceptance criteria as pertains to the instructor. These issues are summarized below.

- The instructor should have a minimum level of expertise in shiphandling. This would ensure that the instructor has the basic skills and knowledge to conduct an effective training program. This requirement should be incorporated as part of the training system acceptance criteria.

- The instructor should exceed a minimum level of expertise in instructional skills. This requirement will be more difficult to specify than that of expertise in shiphandling. Nevertheless, it does represent an important area of instructor characteristics and should be included as part of the training system acceptance criteria. The minimal standards could be based upon (1) some formal or informal training of the instructors in how to conduct an effective training process, (2) experience by the instructors in training, (3) observation by the U.S. Coast Guard, and/or (4) testing of their knowledge of instruction. This is an important aspect of an effective instructor, and it should be included in the standards.

- Training system acceptance criteria should not address the acceptability of an instructor to the trainees. The acceptability of an instructor is a secondary aspect impacting the effectiveness of the training process. If the instructor possesses expertise in both shiphandling and instructional methodologies, he is likely to be acceptable to the trainees. Exceptions will, of course, occur; these can be left to the individual training institution to handle.

- The training program should be appropriately structured with sufficient information and supporting materials provided to the instructor. That is, detailed training objectives should be specified along with detailed topic outlines for each hour of the training program. These should identify the issues to be presented and discussed by the instructor, relevant points to bring out, supporting information, exercises, etc. Much of the information should be obtained in an instructor's guide. Appropriate supporting information should be developed for distribution to the trainees to provide them with additional background information. This type of detailed information would standardize the training program across the instructors and would provide a high probability of successfully achieving the training objectives if followed by the instructor.

- A variety of tools is available for use by the instructor in the conduct of an effective training process. The use of these tools should be taken into account during the

development of the training program structure. The instructor would use the tools to conduct the training process according to the training program's structure, and to ensure that the training objectives are met within the time period available. The tools might consist of feedback displays to present information to the trainee, performance measures to indicate how well the trainee is doing with regard to achievement of various skills, information summaries to be given to the trainees, monitoring capabilities for the instructor to observe details of trainee performance and each exercise, and capabilities to enable the instructor to tailor exercises to the needs of particular students.

The feedback display investigated during this experiment is a limited example of the tools to be used by the instructor during the conduct of training. The discussion of the results found regarding this variable should be looked at, along with the discussion of training assistance technology found in Appendix M. The tools and other supporting training system characteristics augment the instructor qualifications to achieve an effective training process.

The instructional capabilities and tools to support the instructor should be built into each training system. Instructor qualifications (i.e., expertise and acceptability) should be as high as possible, thus ensuring a more effective training process. Minimum instructor qualifications cannot be recommended at this time since the relevant information is not available pertaining to deck officer training. However, as noted above, the two most important characteristics of the instructor are the expertise in shiphandling and the expertise in teaching. From the standpoint of training system acceptance criteria to approve simulator-based training programs, it behooves the U.S. Coast Guard to address these two aspects of expertise as a minimal instructor requirement.

3.10 BRIDGE CONFIGURATION

As discussed in Section 2.2.3, the bridge configuration variable was investigated separately from the main screening experiment. This was accomplished by conducting the master/chief mate training program for a control group (experimental design "B").

It should be noted that due to the difficulty of obtaining qualified instructors for this training experiment, several instructors were prepared for the program. This difficulty is attributable to the nature of CAORF as a research facility, not a training facility. Since CAORF conducts limited

training research, instructors are drawn from a pool of active mariners on an as-needed basis. The schedules of these individuals are not amenable to a 9-week program. In experimental design "A," two instructors were used as part of the experimental design. This allows for the "control" of the instructor effect on the other five variables. In addition, it allows for the measurement of the effect attributable to the instructor variable itself. For experimental design "B," a third instructor was required due to the aformentioned schedule conflicts. Since significant results were expected in favor of the full bridge over the reduced bridge, an instructor who was viewed as CAORF's most qualified instructor was utilized for the reduced bridge group. Therefore, the rationale could be advanced that the uncontrolled instructor variable in experimental design "B" did not contribute to a difference in favor of the full bridge since the more qualified instructor was on the reduced bridge and not the full bridge. However, since few differences were observed between the full bridge group and the reduced bridge group, it should be recognized that this equivalent performance may be the result of the uncontrolled instructor variable in lieu of the effect of the bridge configuration variable.

3.10.1 RESULTS. For the integrated shiphandling training, few statistically significant results were observed for the bridge configuration variable. The group trained on the full bridge had a mean CPA of 234.33 feet in leg 4, whereas the group trained on the reduced bridge had a mean CPA of 157.52 feet (t = 1.566; p ≤ 0.10).

For the emergency shiphandling training, the mean heading deviation from the channel centerline for the group trained on the full bridge increased by 4.8 degrees at data line 40 (leg 2) after training, while the group trained on the reduced bridge had an increase of only 0.1 degree in heading deviation (t = 2.129; p ≤ 0.05) after training.

3.10.2 SCREENING INTERPRETATION. The obtained results show few measurable differences in the effectiveness of the training program as administered on the full bridge versus the reduced bridge. This may be due to a variety of factors, including: (1) simulator fidelity differences between the full bridge and reduced bridge do not have a major impact on the effectiveness of training; (2) the reduced bridge, in the form of a cockpit bridge design, may simply be an effective operational bridge design resulting in better performance; and (3) a well qualified instructor may have effectively compensated for the reduced bridge simulator capabilities. The small number of significant findings, in this case, would appear to warrant further

research since the bridges differed in a wide variety of aspects. That is, further investigation in the screening process is warranted due to the substantial differences between the two bridge designs (e.g., visual scene) and the lack of findings during this experiment. The extent to which each of the three factors noted above contributed to this lack of difference should be delineated in future screening experiments.

3.10.3 EXPERIMENTER'S INTERPRETATION. The project team believes that in this experiment the lack of differences between the effectiveness of training on the full bridge configuration and the reduced bridge configuration was due primarily to the high quality of instruction provided by the particular instructor who was utilized with the reduced bridge group. The researchers believe that this observation re-enforces the primary finding of this screening experiment that the instructor is the most critical element of a simulator-based training system and that he can compensate for some inadequacies in the simulator design. This does not mean that the fidelity of the simulator design should be down-played; only that the qualifications of the instructor should be given appropriate consideration during the design of the simulator-based training system.

It should also be noted that low fidelity bridge configuration may create motivational problems for the students who may question in their own minds whether or not training on the device will transfer to the real world. The instructor utilized in this experiment for the reduced bridge group was apparently able to overcome this problem. However, based on informal discussions with a number of trainees (some of which were in the reduced bridge group; others were not), there appears to be indications that student motivation may be a potential problem when low fidelity bridge environments are employed.

3.11 SUMMARY

The results obtained from this initial experiment of the screening process should be viewed as extremely tentative regarding the relative effectiveness of the different levels of each simulator characteristic investigated. That is, although differences have been found between day and night, 120-degree versus 240-degree fields of view, etc., the results are, at this stage, by no means conclusive. Rather, they indicate that there is reason to continue the research in depth, focused on the highest priority characteristics. Nevertheless, some generalizations can be made regarding each of the characteristics investigated; these are

3-20

discussed elsewhere in this report. An important generalization may be made regarding simulator design characteristics on the basis of the overall set of results obtained by this experiment. The data support the conclusion that the simulator/training device can be cost effectively designed on the basis of a structured subjective training analysis.

The simulators in use today for deck officer training have been designed as a result of two factors: (1) the cost and feasibility of technology to generate the visual scene, etc.; and (2) subjective appreciation for the quantity and quality of information that the mariner uses in achieving his operational objectives. Each simulator in use today differs widely from the other simulators with regard to their specific characteristics. Similarly, they also differ substantially with regard to the training objectives they address during their training programs. It appears that the simulator was designed initially, and then its appropriate application in training was determined. For example, Port Revel obviously lacks the full bridge atmosphere; hence, the staff does not address team training but rather addresses detailed aspects of shiphandling (e.g., bank effects, shallow water effects). By so doing, they train in those areas for which their training facility is strongest and do not attempt to train in those areas for which their facility would likely be inadequate. A similar rationale appears to hold true for most training facilities. That is, the simulator came first; and then, based on its strengths and weaknesses, the training objectives and training programs were developed. Although this approach is backwards from the standpoint that it would be desirable to first define the training objectives, it may nevertheless result in effective training. Ideally, training objectives would be developed initially, and then the remainder of the training system would be developed to most cost effectively meet those objectives. Regardless of the direction in which the resultant training programs were developed, however, either approach would achieve meaningful and effective training.

The design of the currently existing simulators was based largely on subjective evaluation to determine the necessary information, and its requisite level of fidelity, for conducting certain aspects of training. In many cases, the level of fidelity of particular characteristics is purely a function of the technology used in the simulator. Objective data regarding the effectiveness of particular simulator characteristics would likely have the greatest validity in determining the necessary simulator design to meet specific training objectives. That, of course, was the purpose of this Phase 2

experiment. However, the cost in both time and dollars to conduct comprehensive and thorough research in the training area is prohibitive and well beyond the scope of what can reasonably be accomplished. Of substantial importance, therefore, is the observation that the majority of findings regarding the six characteristics investigated were expected and predictable. That is, thorough analysis by training experts would likely have led to the same conclusions reached during this experiment, regarding the characteristics investigated.

The structured training analysis should begin with a task analysis, identification of knowledge and skill requirements, identification of trainee input characteristics, and the resultant development of the specific set of training objectives to be achieved in the training program (see the Phase 1 report for a thorough discussion — Hammell et al., 1980b). If this is accomplished at a highly detailed level, it is likely that the simulator requirements necessary to achieve those training objectives would in fact be adequately identified. For example, if the issue concerns the horizontal field of view that should be specified for a particular simulator (e.g., 120 versus 240 degrees), the analysis would identify those training objectives (e.g., skills) that would and would not be achieved under the two differing fields of view. If one of the objectives was to train the individuals to initiate turns in a restricted waterway on the basis of aids to navigation, the 240-degree field of view would likely be required for much of this training since turn initiation would often occur when a particular aid is abeam of ownship. The data collected in this experiment would support such an analytical approach as likely achieving valid conclusions. (It should be noted that most simulator designs represent a compromise from a training standpoint; some of the desired training objectives are readily achieved while others would not be achieved as cost effectively as desired.) The highly structured, subjective design of the simulator, therefore, is likely to be successful in achieving an acceptable level of cost effectiveness.

The results of this experiment suggest that the highly structured training analysis approach to the design of the training system (i.e., including the training simulator/training device) is appropriate for deck officer training. Resultant simulator designs are likely to be highly effective for the particular sets of training objectives that would be addressed. This is not to suggest that the subjective analysis should replace objective research data. Rather, the training

analysis should be looked at as the tool with which to develop the operational simulators while *objective research* should be conducted continuously, although on a limited basis. The objective research should provide the benchmarks from which the training analysis structure and rationale would be extended. This conclusion is extremely important in that it directly supports the validity of the current approach to simulator design.

Similarly, it is reasonable to assume the U.S. Coast Guard could develop adequate simulator design standards for use as part of training system acceptance criteria, based on a rigorous training analysis. These could be developed for all simulator/training system characteristics, drawing upon mariner and training experts to develop the standards based on their best judgments.

SECTION 4

CONCLUSIONS AND RECOMMENDATIONS

4.1 SIMULATOR AND TRAINING PROGRAM CHARACTERISTICS

The overall purpose of this investigation was to determine the impact of certain simulator/training system character- istics on the effectiveness of deck officer training. Six simulator/training system characteristics were selected for investigation on the basis of their cost and potential impact on training effectiveness (e.g., horizontal field of view). Additionally, two vastly different overall bridge configura- tions were also investigated. This relatively large number of characteristics would require a massive training experiment to thoroughly determine their impact on training effective- ness. A screening process approach, therefore, was decided upon to reduce the overall investigation to a manageable size within the resources available. This approach necessi- tates a series of experiments to determine the most impor- tant characteristics initially and then to narrow the investi- gation to focus only on those important characteristics while ignoring the less important. The initial experient of the screening process is the subject of this report. The pur- pose of the initial experiment is not to exhaustively analyze the characteristics investigated but rather to identify the most important characteristics for more thorough investiga- tion in subsequent experiments. The objective of this experiment, therefore, was to prioritize the six simulator/ training system characteristics for subsequent investigation. Additionally, the experiment sought to make an initial investigation of the impact of vastly different simulator configurations on the effectiveness of training.

The objective of this experiment in the screening process has been met; the priority ranking of the investigated characteristics are presented below. They represent a careful analysis of all the available information, including the statistical tests, observations by the instructors and experimenters, and interpretation by the experimenters. Since the requisite follow-on experiments in the screening process will not be conducted due to a change in the U.S. Coast Guard's timetable for implementing the results, it was determined necessary to go beyond the straightforward objectives of this initial experiment to provide some insight

into the value of the six characteristics investigated and that of the simulator configurations. These additional conclu- sions are based on extended interpretation by the experi- menters. For the most part, the data do not fully support the conclusions reached; rather, the data, together with the observations, etc., suppor* the identification of trends regarding the different characteristics. The trends identified are certainly inconclusive at this time. Furthermore, alter- native explanations are plausible regarding many of the findings. Some of these alternative explanations have been reported in the text. Nevertheless, the extended interpreta- tions by the experimenters are included in this report so as to provide the maximum amount of meaningful informa- tion to the interested user and developer of simulator-based training systems for deck officers.

The results obtained in this experiment are intended to provide information from which decisions will be made by the U.S. Coast Guard and U.S. Maritime Administration. They are not intended to be the final answer regarding particular decisions. Rather, they are intended to act as a primary source of information to assist in making the appropriate decisions. Hence, this information must be carefully analyzed and used with regard to each particular issue as it arises. The information generated during this experiment represents a pioneering effort in the develop- ment of objective data to support the design of simulator- based training systems in the maritime industry.

4.1.1 RANKING OF TRAINING SYSTEM CHARAC- TERISTICS. The prioritization of the simulator/training system characteristics and the bridge configurations was based on analysis of a variety of performance measures and subjective information collected during the experiment. The conclusions reached placed the various characteristics into three categories, indicating their importance for further investigation (see Table 19). The priority I charac- teristics are those which appear to have the greatest poten- tial impact on the effectiveness of training, on the basis of the results of this experiment. These should be investigated in the next experiment of the screening process. The three variables identified (i.e., instructor, horizontal field of view,

TABLE 19. RANKING OF TRAINING SYSTEM CHARACTERISTICS FOR FUTURE RESEARCH (DECREASING PRIORITY CATEGORIES)

Priority I
    Instructor
    Horizontal field of view
    Feedback methodology

Priority II
    Time of day
    Target maneuverability
    Color visual scene

Priority III
    Bridge configuration

and feedback methodology) should be integrated into a single experiment for in-depth investigation to determine their specific impact on the effectiveness of training. Furthermore, these three characteristics should serve as primary candidates for inclusion in design standards for training system acceptance criteria. The appropriate minimum standards regarding each of these would have to be determined from the subsequent experiments of the screening process.

The three characteristics in the priority I category can be further reduced to two characteristics. The instructor and feedback methodology actually, at this stage, represent the nonsimulator related aspects of the training system. That is, the feedback methodology is a tool closely integrated with the instructor; it acts to directly agument the instructor in conducting the training process. Additionally, a variety of other training assistance technology capabilities should be aligned with that of the feedback methodology in augmenting the instructor in conducting his functions. Hence, these two characteristics of the training system are closely integrated and may be considered together in future investigation.

The priority II characteristics were deemed as important in the design of the ship bridge/shiphandling simulator prior to the start of this Phase 2 experiment. The results obtained during this experiment generally support the importance of these characteristics in the design of the simulator/training system. They have been determined, however, to be of relatively less importance that those in the priority I category. Hence, their fuller investigation should follow that of the characteristics in the priority I category.

The bridge configurations have been allocated to a priority III category. As indicated in the discussion section, the results do not show many differences on the impact of training of these vastly different bridge configurations. Due to the wide variety of differing characteristics between the two configurations investigated, the reasons for the results are indeterminate from the data collected in this experiment. Several rationales for these findings are proposed in the discussion. The investigation of different bridge configurations should be undertaken only after a fuller investigation of specific characteristics is made (e.g., priority I and II characteristics) since these characteristics will directly cause the findings that will result in the investigation of the different bridge configurations.

The supporting rationale for the placement of each characteristic in the respective categories is given below:

● **Instructor** — The instructor variable was observed to have the greatest impact on the training effectiveness of those characteristics investigated. The instructor can offset for poorer simulator characteristics as well as degrade a higher quality simulator. The impact of the instructor on the effectiveness of training was at least two to three times as great as that of any of the other characteristics, as observed in this experiment. The instructor is discussed fully in Section 3.9. Future research should address the minimum and recommended qualifications for an instructor to teach deck officers on a bridge simulator, along with the identification of proper instructional techniques to be employed by the instructor. It is imperative that the instructor variable is viewed as more than just the individual since many aspects of the training program and other teaching aids are part of the instructor's scope of influence on training; they directly impact his ability to conduct his functions. The feedback methodology investigated in this experiment is viewed as one aspect of instructor support.

● **Horizontal field of view** — The field of view had a differential impact on the effectiveness of training. In certain situations the lower fidelity field of view (i.e., 120 degrees) was more effective, and in other situations the higher fidelity field of view (i.e., 240 degrees) was more effective. This finding shows that not only is the fidelity of the field of view important, but also the methodological use of field of view as an indirect aid to enhance the training process under certain situations is likewise an important training device design issue.

The concept of achieving greater training effectiveness through the use of lower simulator fidelity is not new, although this is the first time it has been demonstrated via objective data to the experimenters' knowledge. This finding would go counterintuitive to a highly structured analysis of simulator design characteristics on the basis of the operational requirements; rather, it would require careful consideration from the purely training process standpoint to reach such a conclusion. The importance of this finding, coupled with the apparent importance of field of view itself, warrants its inclusion in the highest priority for further research. Findings relevant to the field of view are discussed in Section 3.8.

- **Feedback methodology** — Mixed findings were obtained regarding the feedback methodology. It is included in the priority I category on the basis of the overwhelming strength of the instructor effect. The feedback methodology looked at in this experiment is a very limited example of the training assistance technology that may be integrated into the simulator/training device to augment and assist the instructor conducting an effective training process. This technology is closely integrated with the instructor, in that its purpose is as an aid assisting the instructor in conducting his functions. The effectiveness of the feedback methodology is very closely related to the instructor's ability to properly utilize feedback provided. Additionally, as discussed in Section 3.6, the lack of a stronger effect for the feedback methodology in this experiment may be a result of several factors, including the improper matching of the feedback device with the specific training objectives. Other independent investigations have also noted the potential importance of training aids such as the feedback methodology. Hence, the potential of these training aids, together with their close integration with the instructor which was found to have an overwhelming impact in this experiment, warranted the placement of this characteristic in the priority I category.

- **Time of day** — Results obtained during this experiment, contrary to the opinion expressed by some of the members of the maritime community, suggest that simulator-based training should be conducted under the same ambient lighting conditions as the operational tasks. This finding refutes the theory advanced by some that only nighttime simulator-based training is required since this is the more difficult operational situation. This finding should be further investigated in future research since it casts doubt on the capability of a

nighttime simulator-based training facility to conduct a comprehensive training program. Time of day is discussed fully in Section 3.7.

- **Target maneuverability** — Substantial differences in performance were obtained as a result of the maneuverability of the targets (i.e., canned or freely maneuvering). The specific impact on operational performance, however, is not readily apparent from the results. The interpretation offered by the experimenters is that the freely maneuvering target resulted in better performance than the canned target. However, this interpretation is tenuous and would definitely require considerable further investigation. The preliminary indications are that freely maneuvering targets should be included in those simulators that are training with regard to intership interactions (e.g., rules of the road, meeting situations in restricted waterways). The canned target would likely be adequate for training in those situations where the interaction between vessels is not important (e.g., straightforward navigation in restricted waters). Target maneuverability is discussed in Section 3.4.

- **Color visual scene** — The color visual scene did not appear to improve effectiveness of training for normal shiphandling although it did for emergency shiphandling. Hence, color appears to be relevant under certain circumstances while immaterial under others. Several explanations for the findings are plausible and have been presented in Section 3.5. Further investigation is required since this variable was found to have an impact under certain circumstances.

- **Bridge configuration** — Few differences were found between the effectiveness of training on the full CAORF bridge and the greatly reduced bridge configuration. These two configurations differ widely in their characteristics; hence, the specific cause for the small number of findings cannot be determined. It may be due to several factors, including offsetting design aspects. Simulator/training system characteristics, such as the other six investigated in this experiment, are the factors that directly cause the differences in training effectiveness of different bridge configurations. Since differences in these characteristics have been identified, the lack of differences in the vastly different bridge configurations are likely due to offsetting factors. Hence, the basic information relating to specific characteristics is more important at this stage since it directly delineates the differences in training effectiveness.

4.1.2 RELATED CONCLUSIONS

4.1.2.1 Simulator Design Approach. The findings obtained in this experiment with the possible exception of some of those regarding the field of view, were predictable. That is, a highly structured training analysis conducted by operational and training experts would likely have reached similar conclusions regarding the effectiveness of the various characteristic levels (e.g., day versus night) in achieving the normal and emergency shiphandling training objectives. These data, therefore, support a highly structured although subjective training analysis design approach for the specification of simulator/training device characteristics. This, of course, is the approach that should have been followed in the design of each training device; it is likely that it has been followed to some extent in the design of each device. The results suggest that this type of a highly structured training analysis (e.g., development of tasks, skills and knowledge, training objectives), would result in a cost effective design.

The quasi-subjective design approach should be based on the available subjective information. Hence, objective research should continue to develop the baseline data from which the extended subjective analyses could be initiated. Training system design, therefore, would have objective data as its base, and highly structured subjective analyses for the full development.

4.1.2.2 Simulator Fidelity. The specific simulator characteristics (e.g., field of view) had substantially less of an impact on the effectiveness of training than did the instructor. This indicates that although the simulator characteristics are likely to impact the effectiveness of training, the other nonsimulator aspects of the training program (i.e., the instructor and those aspects that he represents such as the curriculum and support materials) should receive substantial emphasis in the design and accreditation of the simulator-based training programs. It would appear that a minimum level of simulator fidelity for the various characteristics should be required. Little increase in training effectiveness would likely result if the simulator characteristics were at a higher level of fidelity for many of the training objectives. Of course, substantially higher levels of fidelity would be necessary for specific training objectives, although these are likely to be in the minority. Hence, a minimum level of acceptable fidelity should be established, after which the greater emphasis should be placed on the other training program-related aspects of the training system. Likewise, the training system acceptance criteria put into effect by the U.S. Coast Guard should

also establish a minimal set of simulator fidelity standards, along with a set of standards pertaining to the other aspects of the training program (e.g., instructor, curriculum material, and training assistance technology).

Finally, the minimal set of simulator design standards should take into consideration the potential enhanced training effectiveness as a result of lower levels of fidelity which may be possible in certain circumstances. This entails not only identification of the design of particular simulator characteristics from the purely training standpoint (i.e., to focus the student's attention) but also the recognition of the need to compensate for this lower fidelity in other aspects of the training program (e.g., a progressive training program structure to later bring in those aspects that are eliminated with the lower fidelity characteristics).

Although the above conclusions regarding simulator fidelity are somewhat general, they do establish important aspects to be included in the training system acceptance criteria. The structure developed for these criteria should account for those aspects noted above regarding the use of lower fidelity and the importance of the nonsimulator training program aspects of the training system characteristics.

4.1.2.3 Training Assistance Technology. The utilization of advanced concepts of training technology may have potential to greatly enhance the effectiveness of the training process. The strong effect of the instructor observed in this experiment indicates that techniques which improve the instructor's ability to teach the desired skills could have a substantial positive impact on the effectiveness of mariner simulator-based training. Such techniques may include the training methodologies, classroom/simulator mix, use of classroom aids, curriculum characteristics, instructor training, amount and type of feedback, exercise design, and others. It is essential that the cost effective training system be based on careful consideration of the nonsimulator as well as simulator aspects. A brief summary of potential training assistance technology capabilities is provided in Appendix M.

4.1.2.4 Training Effectiveness. The 3-day simulator-based training program developed for this research was found to be an effective means of improving the integrated and emergency shiphandling skills of deck officers. This conclusion does not mean to imply that simulator-based training programs should be configured for 3 days in length. Rather, it shows that a 3-day training program can be effective if properly structured. Of greatest importance, it

indicates the potential effectiveness of simulator-based training for deck officers. The transfer of mariner simulator-based training to the at-sea environment, although assumed, still remains to be established.

4.1.2.5 Port XYZ. The simulated Port XYZ appears to be an acceptable test scenario for future research and training at the master/chief mate level, although it should be validated and appropriately modified to better achieve its goals of evaluating shiphandling skill. This research further established that an effective training program for masters could be designed and implemented within the geographic area contained in Port XYZ. In addition, it should be noted that Port XYZ contains many of the attributes recommended by the SNAME Panel H-10 (Controllability) to be included in a standardized test for evaluating the handling of large tankers entering or departing a representative port (SNAME Panel H-10, 1975).

4.1.2.6 Performance Measures. There appears to be no single performance measure for evaluating shiphandling performance in restricted water scenarios as represented by Port XYZ. The most effective method of evaluating the performance of trainees appears to be through utilization and application of several performance measures. It is expected that a simulator-based training facility with on-going training would be in a position to refine the performance measures utilized to evaluate students on their particular scenarios.

It should be noted that performance evaluation is not considered as a necessary part of the training process. Performance evaluation is necessary to determine the amount of training gain achieved over the training program, and for other evaluations of the trainee population. However, it should not be viewed as a necessary part of an effective training process. The performance measures are necessary for training but only for the presentation of relevant performance-related information to the trainees so as to enable them to learn and understand the relationships between the various parameters they are involved with (e.g., impact of rudder angle on turning circle; impact of range of maneuver on resultant CPA). The performance measures, which are extremely important for training, do not have to provide a single overall indication of performance; furthermore, a particular performance measure may often conflict with other performance measures, as is the case when the deck officer is faced with mixed results for any particular choice of actions. The performance measures, even when conflicting, will provide the trainees with useful information to relate their actions

to particular outcomes in the operational situation. Hence, a variety of performance measures should be used during any training exercise to generate and provide the necessary information to the trainees and thus enable them to learn the relationships of interest and achieve the appropriate skill levels.

4.2 ADDITIONAL RECOMMENDATIONS

The preceding conclusions and recommendations pertain to the specific aspects of this initial experiment of the screening process, and extend conclusions and recommendations derived from it pertaining to the design and use of simulator-based training systems for deck officers. Two major areas of additional recommendations are addressed below, drawing upon the findings of this experiment and the experimenters' interpretations. These address (1) training system acceptance criteria and (2) future research priorities. The training system acceptance criteria for evaluating and approving simulator-based training systems for meeting some partial licensing requirements are currently under development by the U.S. Coast Guard. Several recommendations in this regard are made below. The second set of recommendations below addresses the direction of future research drawing upon the findings of the experiment.

4.2.1 TRAINING SYSTEM ACCEPTANCE CRITERIA

- The particular set of training objectives to be achieved by the training system seeking accreditation should determine the particular simulator acceptance criteria (i.e., minimum design standards and test performance standards) that are used for evaluation. That is, the training system acceptance criteria should be developed either (1) for a particular set of training objectives which relates to a particular license requirement, or (2) in a flexible format so that it can be readily adapted to the particular training objectives that are being addressed by any training facility. In either case, the training system acceptance criteria should be closely tied to the particular set of training objectives that are to be achieved. Minimum training system design and test standards are only relevant with regard to a particular set of training objectives; if the set of training objectives is changed, the minimum design and test standards should also be changed. The remaining recommendations of this subsection pertain to the particular training program used in this experiment, and their principles would apply to other training programs (i.e., other sets of training objectives) as well although not necessarily their specifics.

● Considerable care should be given to the selection of the instructor. This research indicates that the instructor is an extremely important factor relating to the effectiveness of a simulator-based training program. Relatively little is known regarding those qualifications that will make for an effective instructor. It is recommended that the U.S. Cost Guard place emphasis in the training system acceptance criteria on two aspects of instructor qualifications: (1) that he meet minimum standards of expertise in shiphandling (i.e., for shiphandling training), and (2) that he meet minimum standards for expertise in instructional skills. Each of these is considered as a requisite characteristic of an effective instructor. A third aspect, acceptability to the students (e.g., holding an appropriate deck officer's license), is not recommended for inclusion in such standards. Although this aspect may impact the effectiveness of the instructor and the training program, its impact is likely to depend largely on a particular instructor's manner and other indeterminate factors. This aspect should be left to the prerogative of the particular training facility since it should have a greater impact on the acceptability of the training program to the students/deck officers than on the effectiveness of the training program. Additional research is required to more adequately define the attributes of a well qualified instructor.

● A minimum horizontal field of view should be specified. The 120-degree field of view investigated in this experiment did appear to be generally adequate for the training objectives addressed, with several exceptions as noted in the discussion. A large or minimum field of view may be necessary for certain training objectives. Additionally, a smaller minimum field of view may also be acceptable for a subset of the training objectives investigated, or perhaps for all of them; the minimal field of view for the training objectives investigated cannot be determined from the collected data.

● A variable field of view is recommended to provide more cost effective training for normal shiphandling in restricted waters. This recommendation should certainly be allowed although not incorporated as part of minimum simulator design standards. It appears that a variable field of view can increase the effectiveness of training and hence also the cost effectiveness of training. However, it is likely that the large fixed field of view would also be adequate in achieving the training objectives although perhaps not as efficiently (i.e., more training time could be required to achieve the same level of performance). Caution, however, should be exercised

in manipulating the horizontal field of view so as to achieve enhanced training effectiveness. It is recommended that when a reduced field of view is used to focus attention and otherwise enhance the training process, it should be followed by some amount of training under conditions of the higher fidelity level so as to enable adequate transfer to the operational situation.

● A training subsystem, consisting of a variety of training assistance technology such as the feedback techniques employed during this experiment, should be integrated into every simulator/training device. The particular technology that should be integrated for deck officer training depends on many factors, including the specific training objectives, the particular level of trainees, etc. Nevertheless, this capability should be included in the design of every simulator/training device. These capabilities, which improve the instructor's ability to teach the desired skills and knowledge, are expected to have a substantial positive impact on the effectiveness of simulator-based training. Caution, however, should be exercised to tailor the feedback and other assistance techniques to the particular training objectives, etc. It is further recommended that the training system *acceptance criteria address the inclusion of the training subsystem* along with the simulator subsystem in the design of the training device. That is that provision be made to incorporate capabilities to the support of the training process in the simulator/training device. These would consist of the generation of performance-related information and the provision for its feedback to the trainees (e.g., via graphical display means). The training system acceptance criteria should not tightly specify the specifics of these characteristics; they should, however, require their inclusion on a more general specification basis. This would enable the creative development of training assistance capabilities in support of training *rather than tie a training facility to some minimal* standard capabilities. This aspect of the training system is in its infancy and remains to be adequately developed. Nevertheless, it should be recognized and included in minimum design standards.

● Training system acceptance criteria should require that training occur under ambient lighting conditions for which the skills will be used in the at-sea environment. That is, if the training objectives require daylight shiphandling in restricted waters, etc., the minimum design standards should require a daylight visual scene. Likewise, the design standard should require a nighttime

visual scene for training objectives relevant to nighttime shiphandling. The data collected do not fully support this recommendation. The data do form a trend in this direction, and lacking any information to the contrary it appears only prudent to train on the simulator under the operational lighting conditions anticipated at sea. It should be noted, however, that many training objectives may be independent of the ambient lighting conditions; this, of course, depends upon the specific details of the training objectives being addressed.

● It is recommended that an independently maneuverable traffic ship capability be considered for the design of any simulator that will be addressing training objectives regarding traffic ship interactions. Although this research does not firmly establish a basis for this recommendation, it does indicate that the target maneuverability variable is critical for effective training. Until further research identifies situations in which canned traffic ship behavior is acceptable for training, it appears only prudent to recommend the independently maneuverable traffic vessels since they are capable of more accurately modeling the irregular behavior encountered at-sea.

● A black and white visual scene may be acceptable as a minimum standard for many shiphandling training objectives. One exception to this recommendation probably occurs under nighttime conditions when a substantial number of traffic vessels are encountered. Under these conditions, the black and white visual scene may result in less effective training due to the relatively large number of colored lights that would have to be simulated via blink coded patterns. There are undoubtedly other situations in which color may be required for effective training. Caution should be exercised in using a black and white visual scene since color cues have been historically used by the maritime community as a principal source of information for the deck officer.

● Training system acceptance criteria should be based on (1) minimum design standards and (2) test performance standards. It is recommended that the quantity and quality of information currently available pertaining to minimum design standards are inadequate for fully basing the training system acceptance criteria today. The initial training system acceptance criteria, therefore, should be based on generalized principles relating to design standards (i.e., as noted above and others) and heavily dependent upon test performance. Note the test would be to evalute the training facility on the basis

of the performance of a sample of its graduates; the test would not be administered to each graduate to evaluate his individual competence. The initial training system acceptance criteria would thus place the initial heavy emphasis on test performance and some emphasis on general design standards, and would evolve as design and/or test information is developed via research to a more balanced combination of minimum design standards and test performance. The ultimate goal would be to have both the minimum design standards and test performance standards independently provide adequate criteria of training system effectiveness. Their integration into a single set of criteria, therefore, would result in the validity and reliability of training system acceptance criteria exceeding that of either set of standards.

The abov. recommendations address the training system acceptance criteria to be developed by the U.S. Coast Guard to evaluate and approve simulator-based training programs for meeting some deck officer licensing requirements. It should be cautioned that the above recommendations are made to provide information from which the training system acceptance criteria would be developed. The specific issues surrounding the training system acceptance criteria and the use of this data in their development have not been adequately explored during this experiment. Hence, the above conclusions and recommendations should be carefully investigated and evaluated with regard to their specific application in the training system acceptance criteria and their impact on deck officer training facilities. An independent study of these conclusions and recommendations must be made with regard to their specific application to training system acceptance criteria.

4.2.2 FUTURE RESEARCH. The following recommendations pertain to the course of future research based on the findings of this experiment and the developments in the field of mariner training or the course of this investigation. The recommendations are made in two categories according to their priority. The priority I recommendations should be accomplished first, followed by the priority II recommendations.

4.2.2.1 Priority I

● A transfer of training investigation should be conducted to fully evaluate effectiveness of simulator-based training. This investigation would evaluate the transfer of skills and knowledge learned on the simulator to their practice in the operational at-sea situation. A recurrent issue raised by practicing mariners, ship operators, etc.,

4-7

concerns the bottomline effectiveness of simulator-based training. That is, will the deck officers who were trained on a simulator perform better in the at-sea situation? The transfer of training has occasionally, though seldom, been established in other fields (e.g., Smode, 1971). This type of investigation has not been conducted in the maritime industry at any level. It should be conducted to establish a baseline validation of simulator-based training. The transfer of training investigation, although difficult to accomplish, may be achievable in several different ways. This type of investigation should be considered of top priority.

● Specific minimum standards should be determined for each of the characteristics in this experiment. The appropriateness of this minimum standard should then be empirically investigated by comparing effectiveness of training under that condition versus conditions higher and lower than that standard in fidelity. By so doing, the appropriateness of each standard could be objectively evaluated.

● An investigation should be initiated to delineate the characteristics of an effective instructor for mariner simulator-based training. These characteristics should be investigated as a function of trainee characteristics, the training objectives, and other factors of the training situation.

● A research program should be initiated to investigate the effectiveness of alternative training assistance technologies available as an integrated element of the training system (e.g., augmented feedback). This research program should investigate the wide variety of available training assistance characteristics and their effectiveness as a function of other training program and situation factors (e.g., training objectives, exercise conditions).

● An investigation should be initiated to determine those simulator characteristics that could be manipulated to enhance the effectiveness of the training system (e.g., the greater effectiveness under certain circumstances of the lower fidelity 120-degree horizontal field of view, as found during this experiment). The investigation should also delineate other exercise, training objective, situation characteristic, training program structure, and other factors that enable the reduction in simulator fidelity levels to enhance the training process.

● Part-task trainers should be investigated as an adjunct to whole-task simulator training. It is recommended that the investigation seek to delineate the roles of the part-task and whole-task trainers within the overall training system and that both be considered as potentially cost effective complementary elements of the training system.

● The screening process investigation, of which this was the initial experiment, should be continued to fully investigate the identified simulator/training system characteristics. Several of the above recommendations have been made independently of the screening process continuation although they would logically be a part of that continued investigation. It is recommended, as part of this screening process approach, that the simpler and more straightforward experiments of the screening process be conducted. These experiments are likely to generate considerably more meaningful information (i.e., pertaining to training system design characteristics) than did this initial experiment due to their much lower complexity and their ability to fully investigate the respective characteristics.

4.2.2.2 Priority II

● A research program should be initiated to fully investigate the relationship between daytime and nighttime simulator training, and daytime and nighttime ship-handling. The investigation should specifically focus on the differences in deck officer skills for daytime and nighttime operations.

● Additional research is desirable to identify the specific situations, or characteristics of specific situations, in which the independently maneuverable targets would not be necessary, and canned traffic vessels would be acceptable for training.

● Additional research should be conducted to determine the effectiveness of the black and white visual scene for nighttime and daylight training. This research should specifically attempt to identify any interactions between the color/black and white scene and other relevant simulator and situation variables (e.g., number of traffic ships, difficulty of the situation).

APPENDIX A

STUDENT HANDOUT PACKAGE

PROGRAM TRAINING OBJECTIVES

There are four objectives under the categories of integrated shiphandling and emergency shiphandling which will be trained throughout this program:

A. Integrated Shiphandling

    1. Maneuver the vessel through the channel maintaining intended track when either:

        a. The navigational range structures available for various channel legs have a light extinguished, one or both range structures obscured, or one structure missing.

        b. The buoys available for various legs of the channel are extinguished, off position, or missing.

    2. Maneuver the vessel through sharp bends or blind turns into a "Y" channel maintaining intended track and safely avoiding any vessel traffic.

B. Emergency Shiphandling

    1. Maneuver the vessel, maintaining ship control as best as possible, when a rudder failure occurs in any channel leg or turn.

    2. Safely maneuver the vessel when a degradation in the amount of power, or a complete power failure occurs in any channel leg or turn.

C. Overall Objectives

    Safely maneuver the vessel through any leg or turn in the channel when any of the following occur: (1) aid to navigation malfunction, (2) rudder failure, or (3) power failure.

STANDARD WEEK (ONE GROUP)

	7	8	9	10	11	12	13	14	15	16	17	18	REQUIRED SIMULATOR HOURS	
MON				PRETEST								5 HRS	0800–1300
TUE	CLASS NO. 1		SIM NO. 1		L	CLASS NO. 2		SIM NO. 2					4 HRS	0900–1100 AND 1400–1600
WED	CLASS NO. 3		SIM NO. 3		D / L	CLASS NO. 4		SIM NO. 4					4 HRS	0900–1100 AND 1400–1600
THU	CLASS NO. 5		SIM NO. 5		L	CLASS NO. 6		SIM NO. 6					4 HRS	0900–1100 AND 1400–1600
FRI	POST TEST												4 HRS	0700–1100

L ► LUNCH

D ► DEMONSTRATION

A-2

STUDENT HANDOUT

CAORF SIMULATION
BRIDGE

A realistic shipboard environment is achieved in CAORF by means of a full-scale bridge with a complement of actual bridge hardware that can be found on most large contemporary merchant vessels. The flexible design of the bridge facilitates varying the equipment suite and physical arrangement, as desired.

Existing bridge instrumentation consists of:

- Relative motion and true motion radar sets for operating and displaying moving target ships and features such as navigational aids, piers, and shorelines normally found in the open sea, harbors, and docking areas

- Gyro pilot steering control stand which includes the helm unit, steering mode control, heading indicator, rate of turn indicator, rudder order, and rudder angle indicators

- Propulsion console consisting of combined engine order telegraph/throttle control, propulsion plant operating mode control, and rpm indicator

- Course and rudder angle indicator

- Bow thruster control, thruster output indicator, and status light

- Pelorus stands

- Collision avoidance system

- Various displays such as gyro repeater, rudder angle indicator, etc.

- Speed log and ship clock

- Engine order telegraph for one engine

- Engine rpm indicators for one engine

- Fathometer

- Wind speed and direction indicators

- Communications equipment including sound-powered telephone, ship intercom system, single-side-band HF radio, VHF radio, and ship whistle

- Loran C

A-4

CAORF SIMULATION
VISUAL DISPLAY

One of the unique and more extraordinary features of the CAORF simulator is the computer-generated visual imagery which simulates the outside world as seen through the bridge windows. The imagery is projected as a television picture around the bridge on a 60-foot diameter cylindrical projection screen covering a field of view 240 degrees in relative bearing and 24 degrees in elevation.

Detailed presentations in full color are provided of other ships, coastlines, buoys, bridges, buildings, piers, and other significant elements of the environment. The scene also includes ownship's forebody superimposed on the centerline of the screen. The visual scene changes in real time in accurate response to own and other ship maneuvering motions. The system is capable of displaying up to 40 moving ships on the radar from which the closest six moving ships are selected and presented in the visual scene simultaneously.

Other unique characteristics of the visual display include the ability to:

- Simulate restricted visibility conditions by altering the color intensity of an object as a function of the distance of the object from ownship, such that the color of the object approaches the color of fog or haze

- Control the illumination level so that either day or night scenes may be simulated

- Vary the relationship of the generated scene to the watchkeeper's eye height above the waterline for the particular ownship being simulated

- Change the data base to simulate any port in the world

STUDENT HANDOUT

PHASE 2 EXPERIMENTAL SHIPHANDLING PROCEDURES

To obtain experimental uniformity, but still allow each subject to perform in a realistic manner, the following guidelines or "standing order" are set forth:

1. All equipment and components of the bridge (e.g., engine order telegraph, radar, whistles, bow thrusters) may be employed by the subject as he would in an actual at-sea situation.

2. The vessel shall be maneuvered to maintain the centerline of the channel as best as possible.

3. The subject must remain alert for other vessel traffic, breakdowns, or casualties, as they may occur at any time.

4. The Inland Rules of the Road will be followed.

5. The subject will maintain a lookout for traffic, monitor a prudent radar watch, and use the VHF to make all necessary calls and replies.

6. A helmsman on watch will assist the subject when on the simulator for steering only. At no time nor in any way shall they advise or divulge information to the subject, unless failures occur.

7. If an improper maneuver is executed such that it causes the vessel to go out of the channel and into an impossible situation, the scenario will be momentarily stopped. The subject will be given a break while the vessel is repositioned in the center of the channel abeam the position at which it originally left the channel. The scenario will then be resumed.

8. The vessel will be presented at an initial position and aligned on the designated course prior to the start of each training scenario which is approximately 15 minutes in length. Each exercise shall be terminated at a specific position. A critique of each scenario will be conducted upon completion of each run.

9. Due to simulator design, the vessel must maintain a 0.5 knot forward speed at all times.

10. All speeds will be understood to mean through the water.

SPECIFIC INSTRUCTIONS AND INFORMATION FOR THE TEST SCENARIO

1. Ship's maneuvering speed (within the realm of good seamanship) may be maintained while navigating through Wyassup Bay. Once abeam Kingston, the vessel shall not exceed 8 knots. Steerage should be maintained at all times as noted in number 9, above.

2. Throttle control mode will be employed with a 10-second delay for engine order responding.

3. The following conditions exist:

       Time of day: daylight
       Visibility: unlimited
       Wind: 15 knot wind from 210°T, gusting occasionally to 25 knots
       Current: 1½ knots and ebbing
       Tide: no allowance for tide is necessary
       Channel depth: 35 feet+

OWNSHIP CHARACTERISTICS

1. Type: 80,000 dwt tanker with superstructure aft

2. Length overall (LOA): 800 feet

3. Length between perpendiculars (LBP): 763 feet

4. Beam: 125 feet

5. Speed: forward — 0 to 18.5 knots
   astern — 0 to 9 knots

6. Propulsion: Geared steam turbine
   23,000 horsepower
   single screw
   direct pilot house control of throttle and telegraph
   variable throttle

7. Bow thruster: 2000 horsepower

8. Propeller: diameter — 23 feet
   pitch   — 19 feet

9. Bridge: aft to bow   — 675 feet
   forward of stern — 125 feet

10. Maximum rudder: 35 degrees

11. Turning circles: deep water (see Figure A1)
    shallow water (see Figure A2)

For the purposes of this experiment, the vessel has the following characteristics:

1. Loading condition: 70 percent loaded

2. Time to go from full ahead to full astern: 180 seconds

3. Draft: 32 feet with no trim.

4. Freeboard: 20 feet at amidships

5. Height of eye at the bridge: 80 feet

6. Pivot: for turning varies as speed moves forward of midships.

7. Using throttle mode of propulsion only, the following rpm versus speed approximations:

	RPM	Ahead (Knots)	Astern Knots
Stop	0	0.00	0.00
Dead Slow	10	1.55	1.00
Slow	20	3.10	2.00
Half	40	6.20	4.00
Full (Maneuvering)	60	9.30	6.00
Sea Speed	90	—	9.00
Sea Speed	120	18.50	—
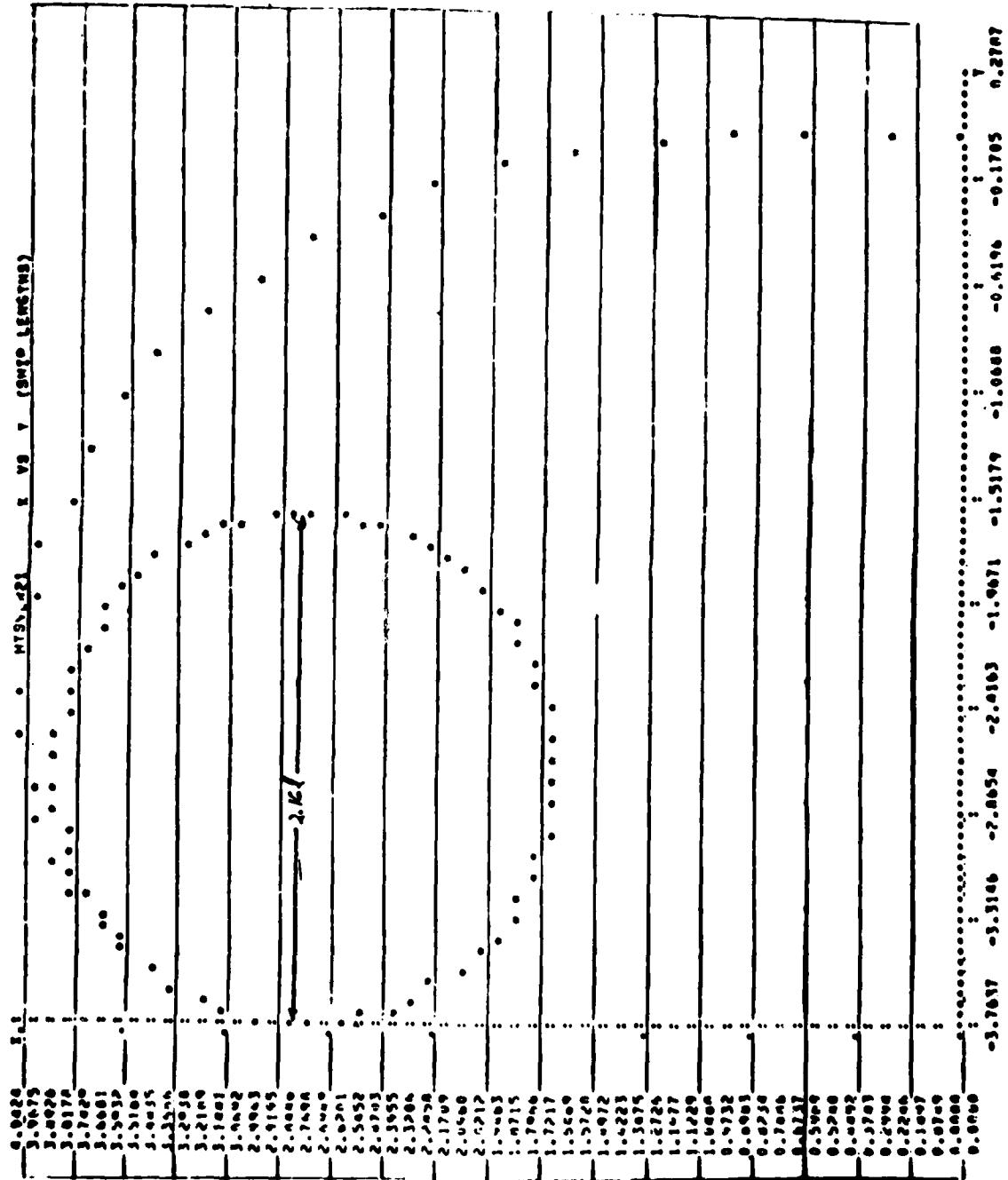
Figure A-1. Deepwater Turning Circle
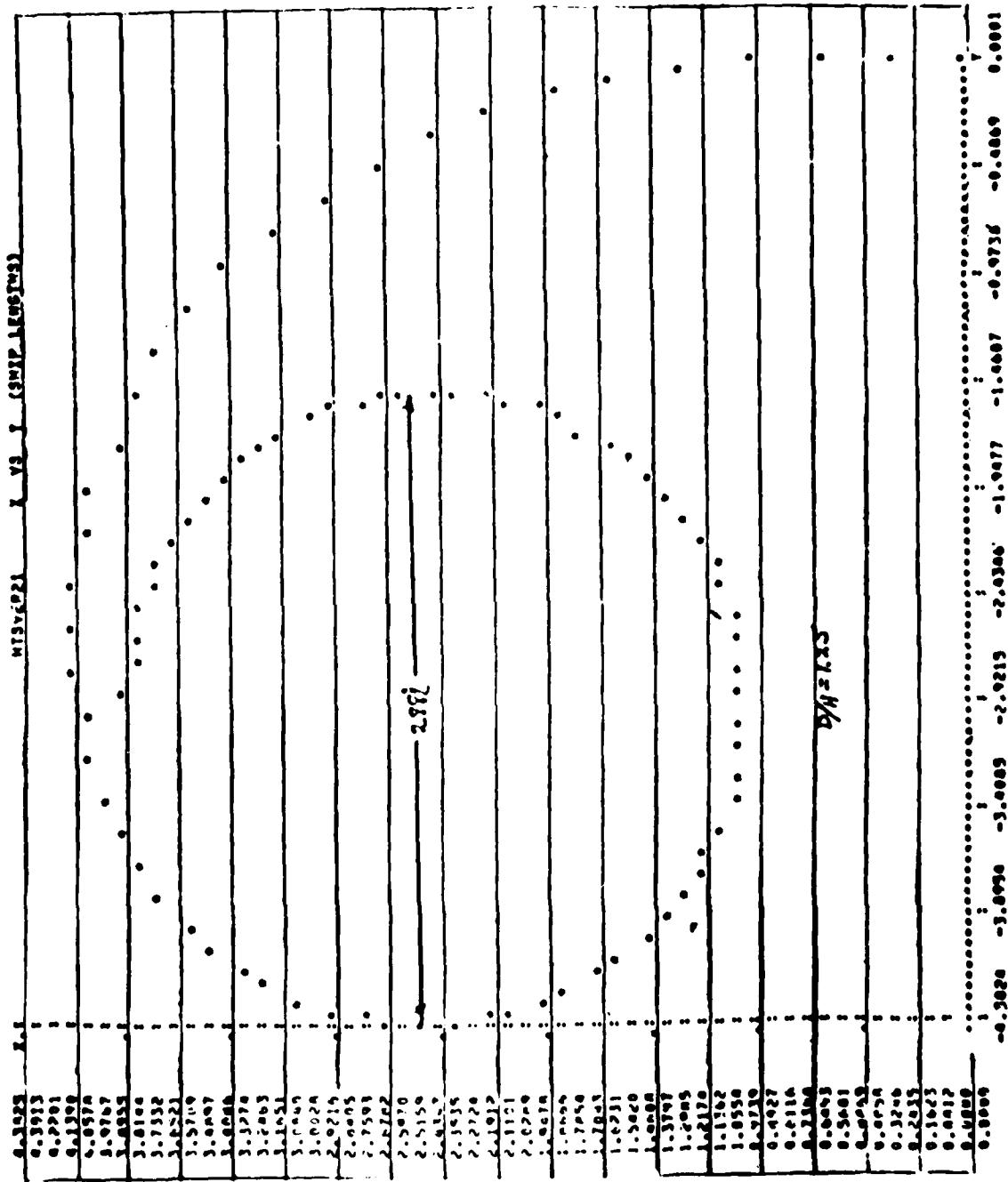
Figure A-2. Shallow Water Turning Circle

A-9

APPENDIX B

TEST DESCRIPTIONS

1. Leg number: 1

2. Turn number: —

3. Description: Misaligned on Ferry Point Range — intend to turn into and transit Gibson's Channel

4. Ownship initial position
   Latitude: 40° 38.0' 44.6" N
   Longitude: 74° 03.0' 52.0" W

5. Ownship initial course: 310°T

6. Ownship initial speed: 8 knots

7. Wind
   Direction (True):   15 knots, gusting to 25 knots from 210°T constant throughout

8. Current: 1.5 knots ebbing — constant throughout

9. Occurrences: Anchored vessel obscures both range lights

10. Traffic: Tug proceeding from Gibson's Channel into Upper Wyassup Bay — crossing situation

11. Communications: See attached

12. End Leg: After passing 74° 05' 23.0" W

NOTE:
   In the pretests and posttests, there will be no realignment of the vessel if it becomes obvious that the vessel is unable to continue. The leg will be stopped and the next leg introduced.

WYASSUP

(OWN SHIP)

START

STOP

KINGSTON

STOP

START

FERRY POINT

C H A N N E L

C O C H

TRAFFIC SIGNAL

B-2

TRAFFIC TYPE   INITIAL CONDITIONS

Target No.	Vessel Type	CAORF Code	Initial		Course	Speed	Remarks
			Position				
1	Containership		40° 39.0' 09.0" N 74° 05.0' 10.0" W				Anchored southwest/ northwest Be sure vessel is high enough to obscure ranges

TRAFFIC MANEUVER TABLE

*Time (min)	Target ___ C S		Target ___ C S		Target ___ C S		Target ___ C S		Target ___ C S		Target ___ C S		Target ___ C S		Target ___ C S	
00	Anchored															

*Time for initial target turn-on

Use additional pages as needed

B-4

COMMUNICATIONS REQUIREMENTS

General Requirements:

1.   Respond only if required

2.   Reply only if indicated for that ship

3.   Respond where indicated using proper procedures

Target No.:_____1_____ Target Type:___Containership_____ Name:_____

Communication Requirement:

   No response

Target No.:_____ Target Type:_____ Name:_____

Communication Requirement:

Target No.:_____ Target Type:_____ Name:_____

Communication Requirement:

Target No.:_____ Target Type:_____ Name:_____

Communication Requirement:

Target No.:_____ Target Type:_____ Name:_____

Communication Requirement:

Target No.:_____ Target Type:_____ Name:_____

Communication Requirement:

TRAFFIC TYPE · INITIAL CONDITIONS

Target No.	Vessel Type	CAORF Code	Initial Position	Course	Speed	Remarks
1	Tug and tow		40° 38.0' 47.0" N 74° 05.0' 50.0" W	075	5	Canned

TRAFFIC MANEUVER TABLE

•Time (min)	Target C	S	Target C	S	Target C	S	Target C	S	Target C	S	Target C	S	Target C	S	Target C	S
00	075	6														
02	060	6			1st Leg											
06	073	6														
09.5	045	6														

•Time for initial target turn-on

Use additional pages as needed

B-7

COMMUNICATIONS REQUIREMENTS

General Requirements:

1. Respond only if required
2. Reply only if indicated for that ship
3. Respond where indicated using proper procedures

Target No.: _____1_____ Target Type: __Tug/Tow_____ Name: __Alice Moran__

Communication Requirement:

    Yes. "Having difficulty. Request I pass ahead of you. Current caught my barge

Target No.: _____1_____ Target Type: __Tug/Tow_____ Name: _____

Communication Requirement:

    Yes. To let tug/tow ahead

Target No.: _____ Target Type: _____ Name: _____

Communication Requirement:

Target No.: _____ Target Type: _____ Name: _____

Communication Requirement:

Target No.: _____ Target Type: _____ Name: _____

Communication Requirement:

Target No.: _____ Target Type: _____ Name: _____

Communication Requirement:

TEST DESCRIPTION

1. Leg number: 2

2. Turn number: —

3. Description: Aligned in Gibson's Channel and intend to proceed toward Fisherman's Point

4. Ownship initial position
    Latitude: 40° 38.0' 49.2" N
    Longitude: 74° 05.0' 41.0" W

5. Ownship initial course: 270°T

6. Ownship initial speed: No change

7. Wind
    Direction (True):   No change

8. Current: No change

9. Occurrences: 4 minute rudder failure commencing when vessel passes 74° 06' 00" west (about 2 minutes into leg)

                                        Note: 10° left rudder to be arbitrarily assigned at this point. Bow thrust available and ship's propulsion OK

10. Traffic: None

11. Finish: End leg 2 minutes after rudder is operational

FERRY POINT

START

STOP

FISHERMEN'S POINT

C H A R L O T T E S B U R G

B-10
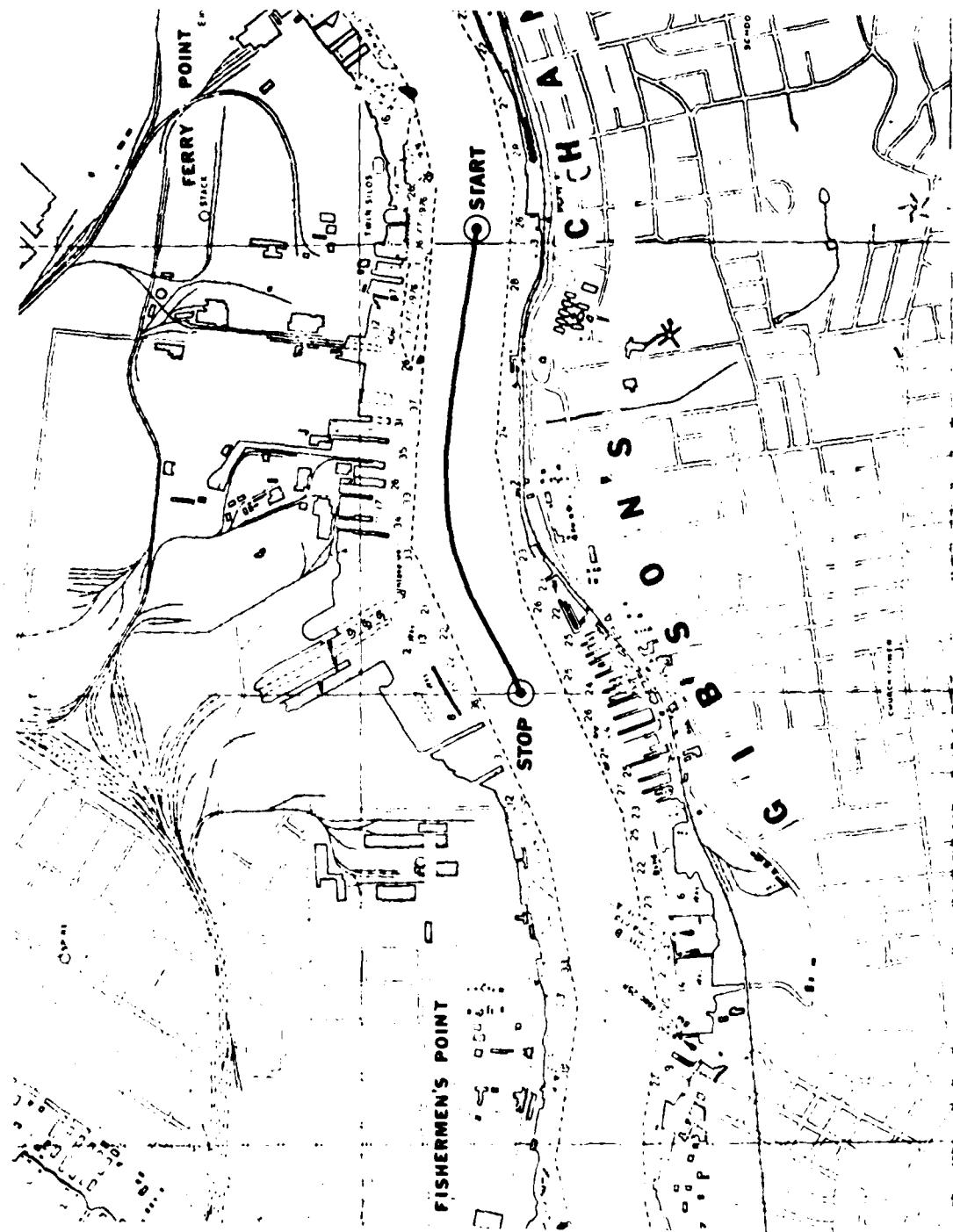
TEST DESCRIPTION

1. Leg number: 3

2. Turn number: —

3. Description: Aligned in Gibson's Channel and intend to proceed toward Fisherman's Point

4. Ownship initial position
    Latitude: 40° 38.0' 47.0" N
    Longitude: 74° 06.0' 50.9" W

5. Ownship initial course: 247°

6. Ownship initial speed: No change

7. Wind
    Direction (True):   No change

8. Current: No change

9. Occurrences: Propulsion plant breakdown of 3 minutes steering and bow thruster available

    Commencement of failure when vessel passes 74° 08.0' 09.2" west (about 2 minutes into leg)

10. Traffic: None

11. Finish: End leg 2 minutes after return of power
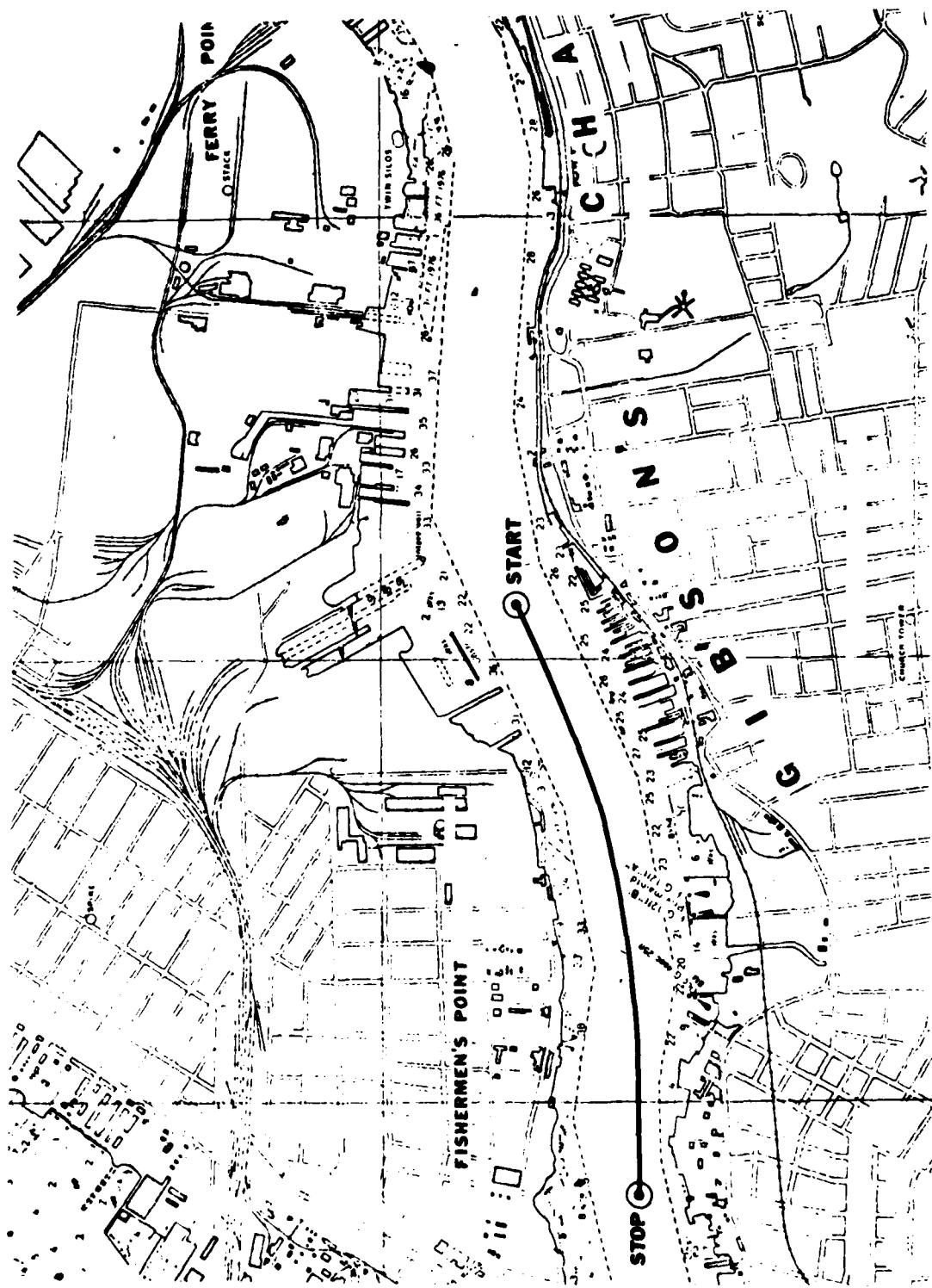
TEST DESCRIPTION

1. Leg number: 4

2. Turn number: —

3. Description: Aligned in Gibson's Channel and intend to proceed Ferry Point into Shellfish Bay

4. Ownship initial position
   Latitude: 40° 38.0′ 33.2″ N
   Longitude: 74° 08.0′ 19.8″ W

5. Ownship initial course: 268°

6. Ownship initial speed: No change

7. Wind
   Direction (True):   No change

8. Current: No change

9. Occurrences: None

10. Traffic: Yes
    A. 1 tug at 40° 38.0′ 27.0″ N
               74° 09.0′ 04.0″ W
       1 tug at 40° 38.0′ 30.0″ N
               74° 09.0′ 10.0″ W
       Both heading north — dead in water

    B. Tug and tow to be shown as own vessel passes 74° 08.0′ 39.0″ W
       Initial position 40° 39.0′ 00.0″ N
                        74° 09.0′ 00.0″ W
       Course 205° proceeding at 6 knots for 1-1/3 minutes and then changes course
          to 142°T (after communication and then south after 74° 09.0′ 00.0″ W
          with ownship)

11. Communications: See attached

12. Finish: End leg when vessel passes 40° 39.0′ 10.0″ N

FISHERMEN'S POINT

START
(OWN SHIP)

START
(TK/AETK
SHIP)

STOP

STOP

Blackfish Island
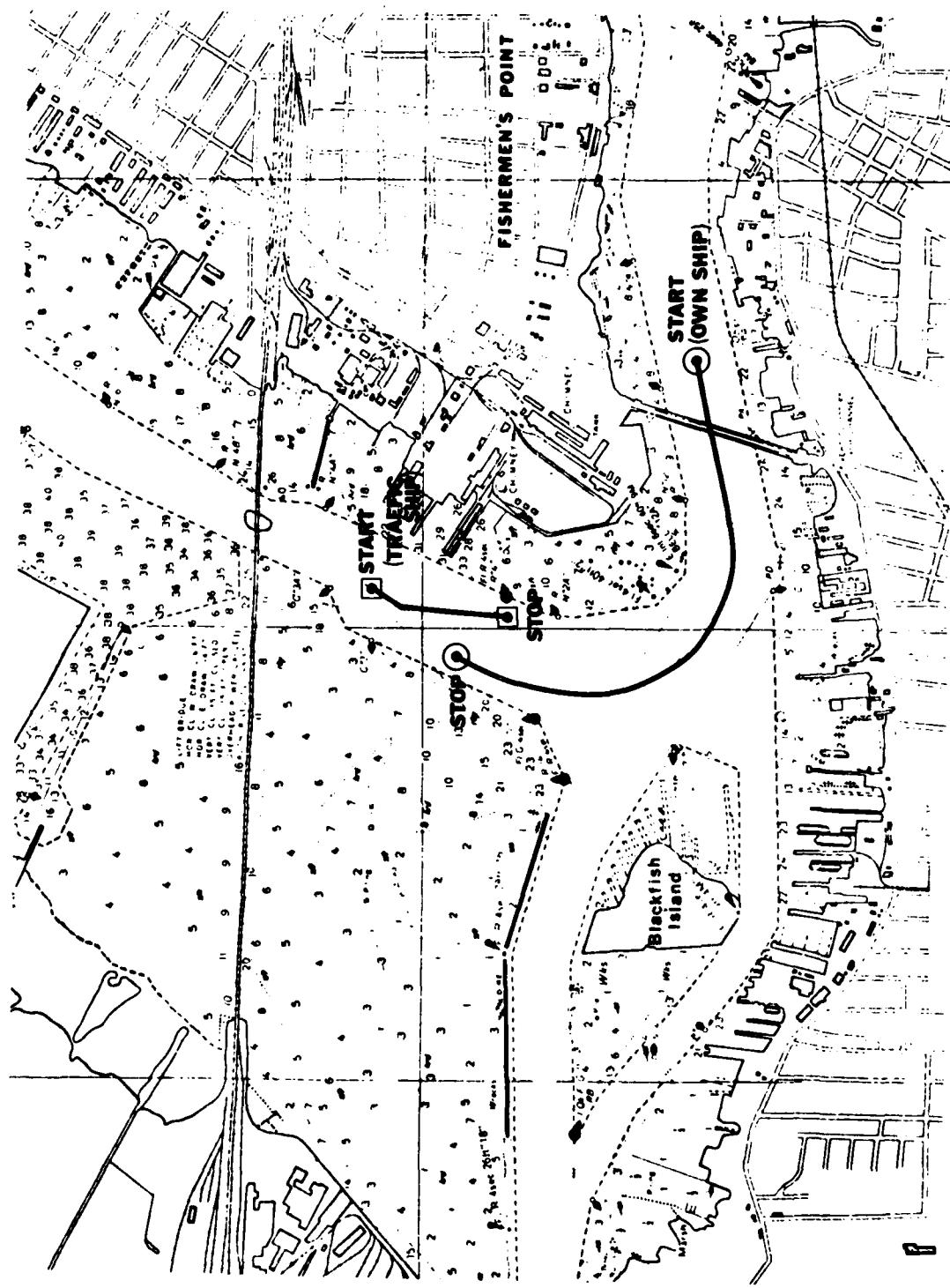
Target No.	Vessel Type	CAORF Code	Initial			Remarks
			Position	Course	Speed	
1	Tug		40° 38.0' 27.0" N 74° 09.0' 04.0" W	Heading north	Dead in water	
2	Tug		40° 38.0' 30.0" N 74° 09.0' 10.0" W	Heading north	Dead in water	

COMMUNICATIONS REQUIREMENTS

General Requirements:

1. Respond only if required
2. Reply only if indicated for that ship
3. Respond where indicated using proper procedures

Target No.: ____1____ Target Type: __Tug_____ Name: _Judy Magallester_

Communication Requirement: ·

   Yes. "Waiting for an inbound sea land vessel"

Target No.: ____2____ Target Type: __Tug_____ Name: _Cynthia Moran_

Communication Requirement:

   Yes. "Waiting for an inbound sea land vessel"

Target No.: _____ Target Type: _____ Name: _____

Communication Requirement:

Target No.: _____ Target Type: _____ Name: _____

Communication Requirement:

Target No.: _____ Target Type: _____ Name: _____

Communication Requirement:

Target No.: _____ Target Type: _____ Name: _____

Communication Requirement:

TRAFFIC MANEUVER TABLE

*Time (min)	Target		Target		Target		Target		Target		Target		Target		Target	
	C	S	C	S	C	S	C	S	C	S	C	S	C	S	C	S
00	205	6														
01.3	142	6														
02.3	170	6									4th Leg					

*Time for initial target turn-on

Use additional pages as needed

COMMUNICATIONS REQUIREMENTS

General Requirements:

1.   Respond only if required
2.   Reply only if indicated for that ship
3.   Respond where indicated using proper procedures

Target No.:_____1_____Target Type:___Tug and Tow_____ Name:_____

Communication Requirement:

   No response

Target No.:_____Target Type:_____ Name:_____

Communication Requirement:

Target No.:_____Target Type:_____ Name:_____

Communication Requirement:

Target No.:_____Target Type:_____ Name:_____

Communication Requirement:

Target No.:_____Target Type:_____Name:_____

Communication Requirement:

Target No.:_____Target Type:_____ Name:_____

Communication Requirement:

APPENDIX C

DATA COLLECTION SHEETS

BIOGRAPHICAL DATA SHEET

(PLEASE PRINT OR TYPE)   DATE:

NAME: AGE:

ADDRESS: HOME

                OFFICE

PHONE NO: HOME

                OFFICE

TRAINING/EDUCATION:

U.S. Merchant Marine Academy   Year Graduated

State Merchant Marine Academy   Year Graduated
                                (specify school)

University   Degree   Year Graduated
        (specify name of college)

Military   Years Served
                                (specify service)

Other
                                (name and duration of course or school)

LICENSED DECK OFFICER EXPERIENCE: YRS

CURRENT AND PAST LICENSES:
(Note pilotage, if applicable)   ISSUE NO.   ISSUE DATE

        Pilot

        Master

        Chief Officer

        Second Officer

        Third Officer

BIOGRAPHICAL DATA SHEET (Continued)

RADAR OBSERVER ENDORSEMENT:

    Yes

    No

TONNAGE RATING:

    Unlimited

    Limited   (specify tonnage)

MILITARY EXPERIENCE AS OFFICER OF THE DECK UNDERWAY:

    U.S. NAVY   YRS.

    U.S. COAST GUARD   YRS.

    OTHER   YRS.

                                (specify)

SEA SERVICE RECORD*

DATES FROM	TO	POSITION	WATCH STANDER**	VESSEL NAME	VESSEL TYPE	GROSS TONS	DWT	PROPULSION HP	TYPE	MAX. SPEED

*Indicate time worked ashore, if applicable

**Indicate YES or NO

If any information is unknown, indicate by "UNK"

C-3

1.0

1.1

1.25 1.4 1.6

4.5 2.8 2.5
3.2 2.2
3.6
4.0 2.0

1.8

MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963 A
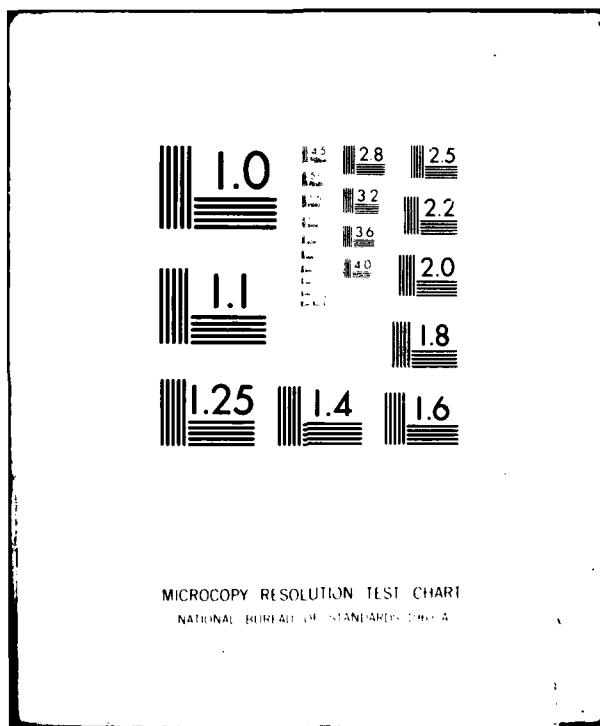
HUMAN FACTORS DATA SHEET

TEST SUBJECT NO.: _____
HELMSMAN: _____
MATE: _____

GROUP NO: _____
P.B. TAPE NO.: _____
SCENARIO NO.: _____

DATE: _____
DATA TAKEN BY: _____
SUBJECT NAME: _____

AID TO NAV.: _____
FAILURE MODE: _____

BRIDGETIME	RPM	HELM ORDER														BOW THRUSTER		FAILURE		DETECTION OF A/N		COMMUNICATIONS				
		HARD	LEFT				MID-SHIP	RIGHT							HARD	R	L	TIME OF NOTIFICATION	TIME OF RESPONSE	YES	NO	SECURITY CALL	EMERGENCY CALL	REQUEST INTENTION	REQUEST POSITION	STATE INTENTION
			35	30	25	20	15	10	5	5	10	15	20	25	30	35											

APPENDIX D

SIMULATOR MODIFICATIONS AND EVALUATION
FOR THE CHIEF MATE EXPERIMENT

Simulator Modifications and Evaluation

In the preparation of any experiment at CAORF finite requirements are generated in order to accomplish the experiment goals. When these requirements fall outside the present capabilities of the CAORF system, either the requirements or CAORF is modified. In the case of the Training and Certification Program all needs were met by existing CAORF capabilities except the following needs which were met by modifying CAORF:

- Black and white visual scene

- Reduced field of view

- Excessive turnaround time between runs

- Track plots used as training aids

- Reduced bridge

1. To satisfy the requirements of the Training and Certification Program Part 2, it was necessary to accomplish the simulation with objects painted black and white in varying shades as opposed to the normal full color display. This was readily accomplished by redefining the colors contained in a color file and their intensity to yield the various gradations between black and white.

   Colored lights were redefined to exhibit distinct white flashing patterns in lieu of red and green. This applied to traffic ship port and starboard lights. The starboard light (green) was a Morse "N" (2 sec flash, off ½ sec, flash, off ½ sec). The port light was a Morse "H" (4 flashes at 1 sec intervals, off 4 sec).

2. A further alteration of the CAORF simulation was to reduce the Field of View. The normal Field of View is 240 degrees. This was reduced to 120 degrees by turning off the two extreme projectors which together accounted for 96 degrees. The remaining 24 degrees were obtained by masking out a 12 degree portion of the projected image on either side to achieve a 120 degree field of view.

3. The Training and Certification Program consisted of 30 scenarios which were contained in 16 System Setup Tapes (SSTs). These were required to implement relatively short training scenarios. As a consequence, considerable simulation time was consumed in initializing and reinitializing the system. To minimize time between runs, recommendations were made to implement SST changes. Normally these changes were done manually via keyboard and would take about 15 minutes for the Training and Certification Program. Utilizing cards to input the SST data approximately 5-7 minutes was required. Utilizing the SST data changes when stored in disc files took 3-4 minutes. Besides saving time, a gain was made in that SST changes were entered more reliably.

4. Training was augmented with trackplots available for post run examination as discussed in Section 4. Difficulty was experienced in obtaining these trackplots immediately after the simulator exercise. The mean leg time experienced was

approximately 15 minutes. To circumvent this problem, the programming to generate plots in parallel with the simulation was implemented and would be available for future programs with adequate lead time to incorporate individual requirements.

5. A variable in the experiment was a reduced fidelity bridge. To meet this requirement the reduced bridge, shown in the attached drawing, was fabricated and installed on the floor of the CAORF theater. This bridge was a wood-framed module (7 ft by 9 ft) with five 25-inch TV monitors mounted in the windows of the bridge. These monitors provided a black and white reproduction of the CAORF visual scene similar to that presented above the human factors station. The following equipment was provided in the bridge:

- Raytheon 16-inch radar display

- RPM order indicator

- RPM indicator

- Speed indicator

- Wind direction indicator

- Wind speed indicator

- Heading indicator

- Rudder order indicator

- Rudder angle indicator

- Rate of turn indicator

Communications for the "reduced" bridge were provided in the following manner: the mate at the bridge main engine control panel and the helmsman were positioned on the full CAORF bridge and instructed to carry out all orders given by the test subject via the installed intercom system.

APPENDIX E

SUBJECT DEMOGRAPHIC SURVEY

The following demographic survey of the subjects involved in this experiment provides an overview of their areas of residence, age range, years of sea experience, union affiliation, their education and the license held. It should be noted that the subject distribution across each of these three categories was not predetermined. In addition, no significance of subject variability is implied for this analysis.

AREAS OF RESIDENCE

For purposes of this analysis, the country was divided into the following four geographic areas: (1) New England, (2) mid-Atlantic states, (3) gulf coast and the southeastern states, and (4) the west coast. As can be seen from Figure E-1, 50 percent of the subjects were from the mid-Atlantic states, 19.2 percent resided in New England, 17.3 percent were from the west coast, and 13.5 percent resided in gulf coast and southeastern states.
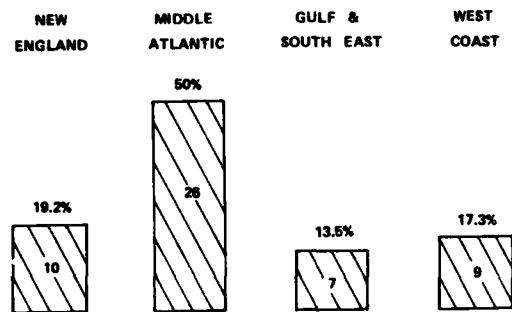
Figure E-1. Areas of Residence

AGE OF SUBJECTS

The subjects range in age from 25 to 70 years. Figure E-2, divides this range into eight segments (1) 25-29, (2) 30-34, (3) 35-39, (4) 40-44, (5) 45-49, (6) 50-54, (7) 55-59, (8) 60+. The average age of subjects was 43.13.
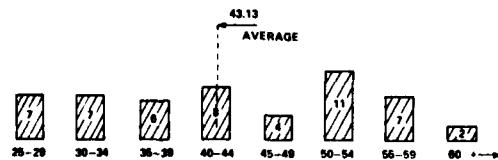
Figure E-2. Ages of Subjects

E-1

YEARS OF SEA EXPERIENCE

The number of years at sea ranges from 3 to 36. Figure E-3 divides this range into 8 segments: (1) 0-4, (2) 5-10, (3) 11-14, (4) 15-19, (5) 20-24, (6) 25-29, (7) 30-34, (8) 35+. The average years of sea experience for the subjects was 15.48 years.
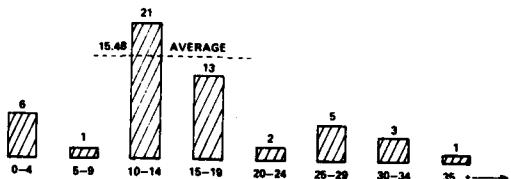
Figure E-3.   Years of Sea Experience

UNION AFFILIATION

The union affiliations of the students who participated in the training program were as follows:

MMP	58%
AMO/MEBA-2	25%
BMO	6%
Other	11%
	100%

EDUCATION

Seventy-one percent of the subjects were graduates of a Maritime Academy or other formal training facility.

LICENSE

Seventy-six percent of the subjects held Masters' Licenses, 22 percent held Chief Mates' licenses, and two percent had a second mate's certificate.

APPENDIX F

ALTERNATIVE DESIGN METHODOLOGIES

Several methodological training research approaches may be followed to accomplish the goals outlined in the introduction of this report. The following is a comparison of the three alternative design methodologies considered. The traditional method for evaluating the effectiveness of differing variables upon training has been the completely balanced design. Employment of this method allows for the evaluation of variables thought to affect training independently of all other variables incorporated in the design. Both main and interaction effects can be completely evaluated in terms of their influence upon the acquisition of to-be-trained skills. Such designs, however, are limited in that many sets of treatment conditions must be run with a different group of subjects for each set. For example, if six variables were to be tested, with two levels per variable, a total of 64 groups of subjects would be required; each group would receive a complete training program. This method is obviously cumbersome with regards to the large number of variables that have been presented for possible consideration in the Phase 2 program. The potential payoff, in terms of information gained, is low relative to the time and cost requirements for such exhaustive experimentation.

A comparison of several complete simulator design configurations is a second investigative approach that examines the macro level of the problem. For example, three well-defined levels of simulator sophistication can be incorporated into an experiment to determine the training effectiveness relative to high, medium, and low levels of simulation fidelity. A complete simulator design would be configured for each level. Such an approach, however, would preclude any analysis or identification of individual simulator characteristics (e.g., field of view). Consequently, the results would not be generalizable to other simulators but would be fixed to those three well-defined representative simulators incorporated into the investigation.

A third methodological approach known as the "screening process" (Simon, 1973) involves the use of experimental techniques categorically referred to as fractional factorial designs. These designs permit the evaluation of a large number of treatments, but employ only a fraction of the total possible number of treatment combinations. These designs also lend themselves to sequential research program — in which flexibility in the pursuit of promising lines of investigation is essential. Furthermore, if it is deemed necessary to add to the design, the flexibility of these fractional factorial designs allows for such additions. This technique permits the systematic investigation of a large number of variables. The early stages of investigation generate limited, and incomplete, information about all the variables. These findings permit the insignificant variables to be eliminated from further investigation, permitting the later stages of investigation to focus more deeply on only the several important variables, and thus generate more complete data on these.

This approach provided the framework used for the Phase 2 program — that of allowing for the integration of many experiments into a consistent structure. The output from such a framework would consist of a series of prediction equations to evaluate the effectiveness of differing levels of training and simulator variables on overall training effectiveness. Such a systematic approach has many advantages, including:

a.   *A consistent methodology that allows for the simultaneous assessment of a great number of variables.*

b.   *The development of a standardized training performance data base relative to this large number of variables, that can be built over time.*

c.   *A timely, efficient process for eliminating variables that have little impact upon the training process, as the investigation proceeds into greater depth concerning the more significant variables.*

d.   *The timely assessment of those variables that make the greatest contribution to the training process via a structured systematic investigative process.*

F-1

There are, however, disadvantages to this approach. The interpretation of the statistical analysis is complicated by the fact that treatments are confounded with higher order interactions in the early stages of investigation. However, if variables are selected carefully so that interactions are assumed to be negligible the confounding of main effects and higher order inter-actions are of little concern. This approach also requires that all or most treatments have the same number of levels and, in that respect, are restrictive. The layout and computational procedures are more complex for fractional factorial designs than for the traditional completely balanced designs. Finally, a sequential series of experiments is usually necessary to achieve conclusive results.

Consideration of both the advantages and disadvantages of the alternative approaches clearly favored the use of the fractional factorial design in order to test a large number of variables in a reasonable time frame.

APPENDIX G

SAMPLE INSTRUCTOR'S GUIDE

UNIT 3
EXERCISE OBJECTIVES

CLASSROOM EXERCISE 5

Determine the possible compensatory procedures to employ in leg 3 and turn 1 of the gaming area when given the following types of rudder failures under the conditions of no wind, no current, and no traffic:

Rudder jammed:
- Amidships (leg 3)
- 15°L (leg 3)
- 15°L (turn 1)

SIMULATOR EXERCISE 5

Maneuver the vessel maintaining ship control as best as possible under the conditions of no wind, no current, and no traffic, given the following types of rudder failures:

Rudder jammed:
- Amidships (leg 3)
- 15°L (leg 3)
- 15°L (turn 1)

CLASSROOM EXERCISE 6

Determine the possible compensatory procedures to employ when given the following types of rudder failures in the presence of the associated environmental conditions:

Rudder jammed:
- Amidships (leg 3, 50 knot wind (from 270°T), 3 knot flood current)
- 15°L (leg 3, 50 knot wind (from 045°T), 3 knot flood current)
- 15°L (turn 1, 30 knot wind (from 270°T), 3 knot flood current)

SIMULATOR EXERCISE 6

Maneuver the vessel, maintaining ship control as best as possible, when the following types of rudder failures occur in the presence of the associated environmental conditions:

Rudder jammed:
- Amidships (leg 3, 50 knot wind (from 270°T), 3 knot flood current)
- 15°L (leg 3, 50 knot wind (from 045°T), 3 knot flood current)
- 15°L (turn 1, 30 knot wind (from 270°T), 3 knot flood current)

SAMPLE INSTRUCTOR'S GUIDE

UNIT 3
CLASSROOM EXERCISE 6

METHODOLOGY

Feedback (postproblem critique) to simulator exercise 5
Presentation of classroom exercise 6 issues in preparation for simulator exercise 6
Positive guidance
Seminar type discussion for both feedback and classroom exercise 6 issues

MATERIALS

1.   Actual track plots of scenarios 15 through 17 completed in simulator exercise 5.

2.   Transparencies of leg 3 and turn 1 (see classroom exercise 5, materials -1) of the gaming area replicated from the chart of Upper New York Harbor showing own ship's intended track.

3.   Transparency of turn 1 showing own ship with arrows indicating the presence of following forces on own ship:

a.   30 knot wind from 270°T
b.   3 knot flood current

4.   Transparency of leg 3 showing own ship with arrows indicating the presence of the following forces on own ship:

a.   50 knot wind from 270°T
b.   3 knot flood current

5.   Transparency of leg 3 showing own ship with arrows indicating the presence of the following forces on own ship:

a.   50 knot wind from 045°T
b.   3 knot flood current

6.   Possible track plots showing own ship's course if no compensatory procedures were employed when own ship encountered the following types of rudder failures in the presence of the associated environmental conditions:

Rudder jammed:

● Amidships (leg 3, 50 knot wind from 270°T, 3 knot flood current, own ship's position when failure begins:
40° 38.0′ 48.0″ N/74° 06.0′ 00.0″ W — failure lasts 4 minutes)

● 15°L (leg 3, 50 knot wind from 045°T, 3 knot flood current, own ship's position when failure begins:
40° 38.0′ 50.0″ N/74° 06.0′ 27.0″ W — failure lasts 1 minute)

● 15°L (turn 1, 30 knot wind from 270°T, 3 knot flood current, own ship's position when failure begins:
40° 39.0′ 04.0″ N/74° 04.0′ 54.0″ W — failure lasts 2 minutes)

7.   Possible track plots (as described in 6 above) when compensatory procedures were employed. Identify each procedure with its associated track plot.

8. Handout

Description of the effects of a rudder jammed (a) amidships in leg 3 with a 50 knot wind from 270°T and a 3 knot flood current, (b) 15°L in leg 3 with a 50 knot wind from 045°T and a 3 knot flood current, (c) 15°L in turn 1 with a 30 knot wind from 270°T and a 3 knot flood current, and the possible compensatory procedures which could be employed when each of the above situations occur.

FEEDBACK (POSTPROBLEM CRITIQUE) (10 MINUTES)

1. Critique, using the actual track plots of scenarios 15 through 17, or verbal discussion only, the successful as well as unsuccessful strategies employed to compensate for the various types of rudder failures inflicted upon own ship.

2. Discuss the reasons for the success or failure of the different strategies employed.

PRESENTATION (15 MINUTES)

1. Using the transparencies, review the layout of leg 3 and turn 1 of the gaming area.

2. Indicate own ship's intended track through each of these two segments.

POSITIVE GUIDANCE AND SEMINAR TYPE DISCUSSION (35 MINUTES)

1. Using the transparencies as aids (see Materials 3-5), discuss the effect of wind and current on own ship while in turn 1 and leg 3.

2. Discuss, using the devised track plots (see Materials - 6), the effects that the following types of rudder failures, in the presence of the associated wind and current forces, would have on own ship if no compensatory procedures were employed:

Rudder jammed:

- Amidships (leg 3, 50 knot wind from 270°T, 3 knot flood current, failure lasts 4 minutes)

- 15°L (leg 3, 50 knot wind from 045°T, 3 knot flood current, failure lasts 1 minute)

- 15°L (turn 1, 30 knot wind from 270°T, 3 knot flood current, failure lasts 2 minutes)

3. Discuss and illustrate (using the track plots from Materials - 7), the possible compensatory procedures that could be employed when given the situations described in 2 above (e.g., judicious use of astern power to bring the vessel to a stop, but leaving a slight headway to help compensate or take advantage of the effects of wind and current).

4. Discuss the effect excess speed (12 knots) would have on own ship while trying to compensate for a rudder jammed amidships or 15°L in leg 3. Discuss various alternate speeds which would be more suitable for handling this type of emergency situation.

ASSOCIATED SCENARIOS

Scenarios 18 through 20

G-3

SAMPLES OF VISUAL AIDS
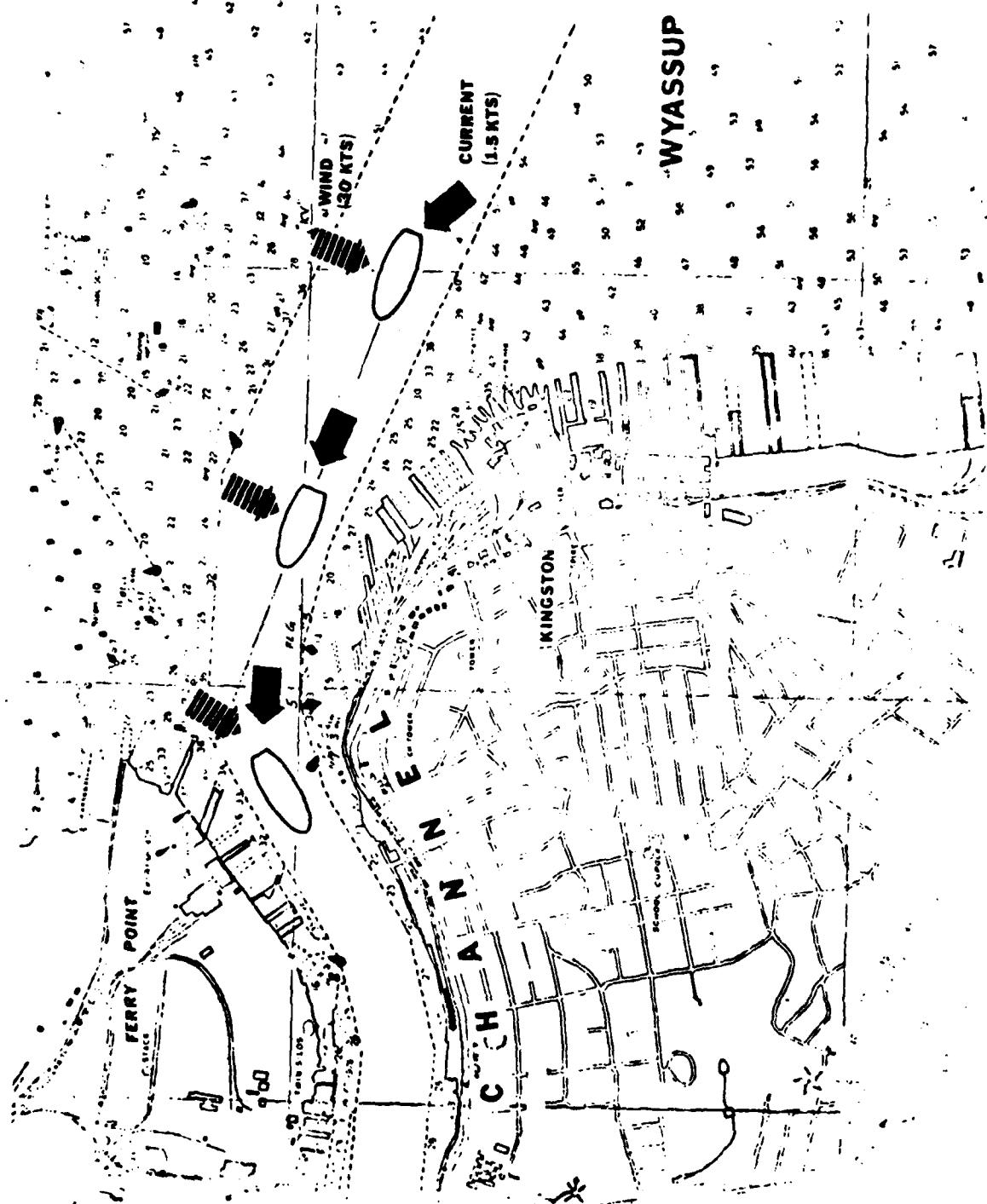
SHIP DYNAMICS PROGRAM - NMRC/KINGS POINT   HARBOR: PORT XYZ
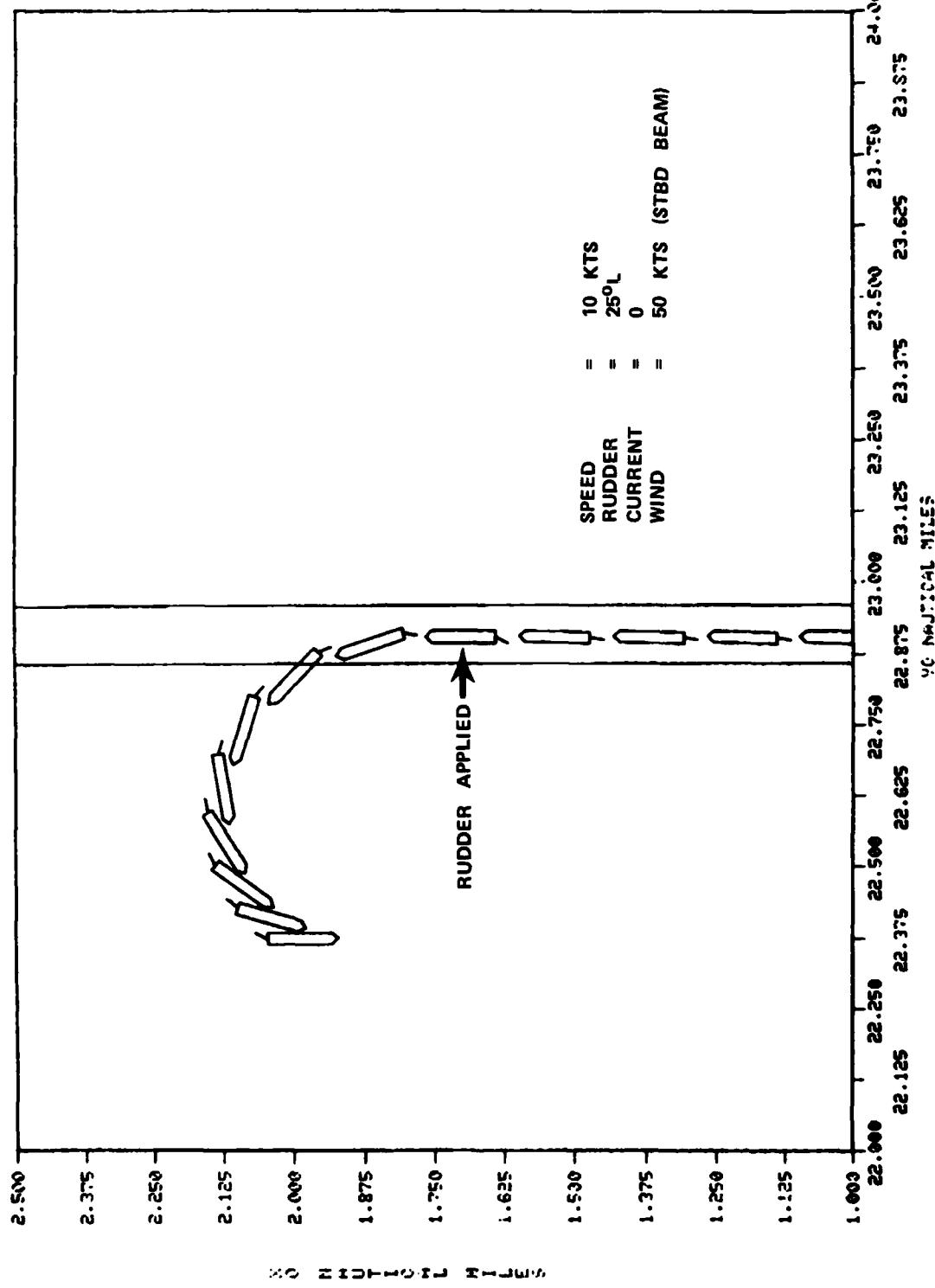
SPEED   = 10 KTS
RUDDER = 25°L
CURRENT = 0
WIND   = 50 KTS (STBD BEAM)

RUDDER APPLIED

H-2

SHIP DYNAMICS PROGRAM - NMRC/KINGS POINT   HARBOR: PORT XYZ

SPEED	=	10 KTS
RUDDER	=	15°L
CURRENT	=	0
WIND	=	50 KTS (STBD BEAM)

RUDDER APPLIED

YO NAUTICAL MILES

XO NAUTICAL MILES

H-3

SHIP DYNAMICS PROGRAM - NMRC/KINGS POINT   HARBOR: PORT XYZ

SPEED   = 10 KTS
RUDDER   = 10°L
CURRENT = 0
WIND   = 50 KTS (STBD BEAM)

RUDDER APPLIED

YO NAUTICAL MILES

XO NAUTICAL MILES

H-4
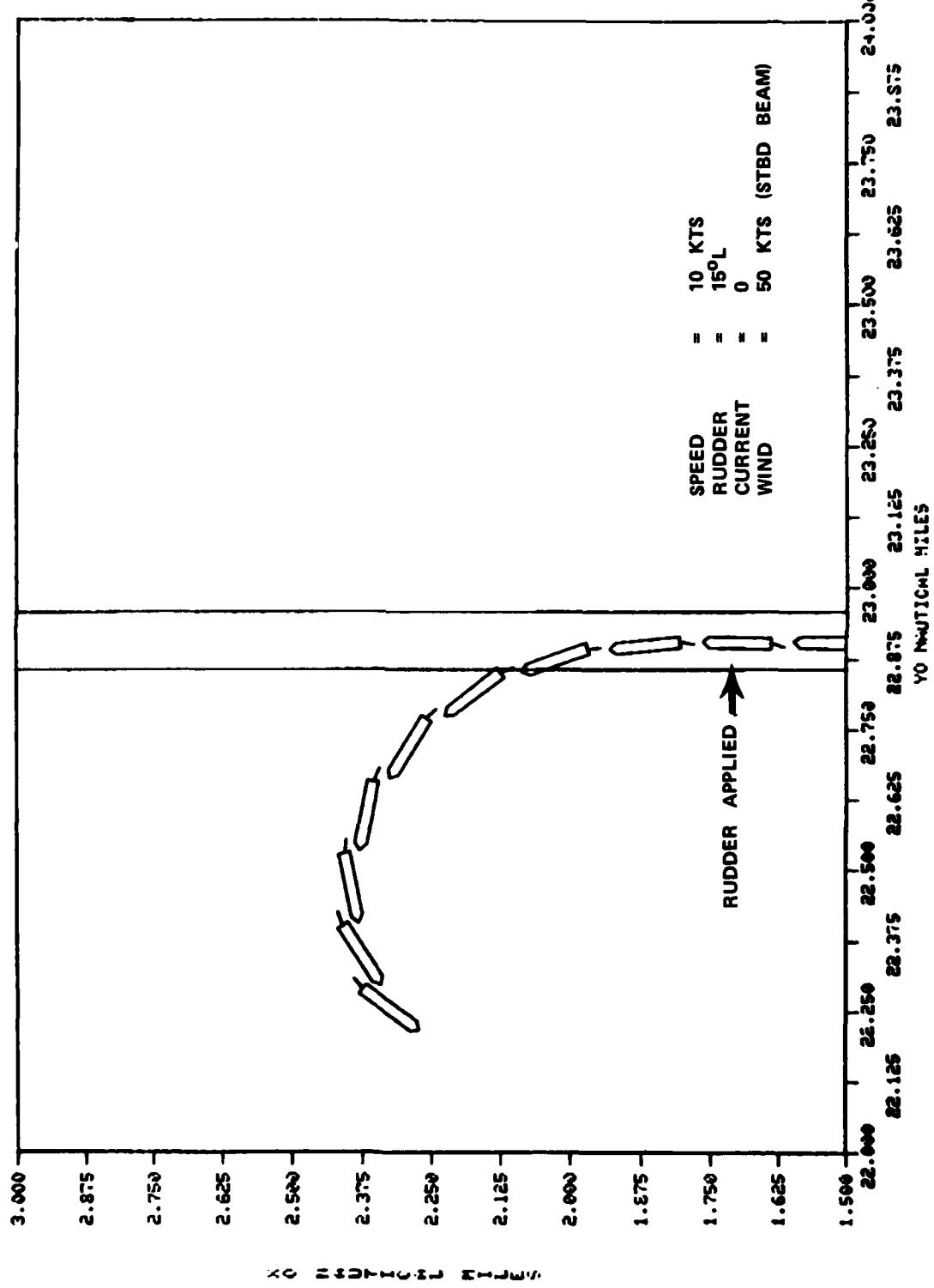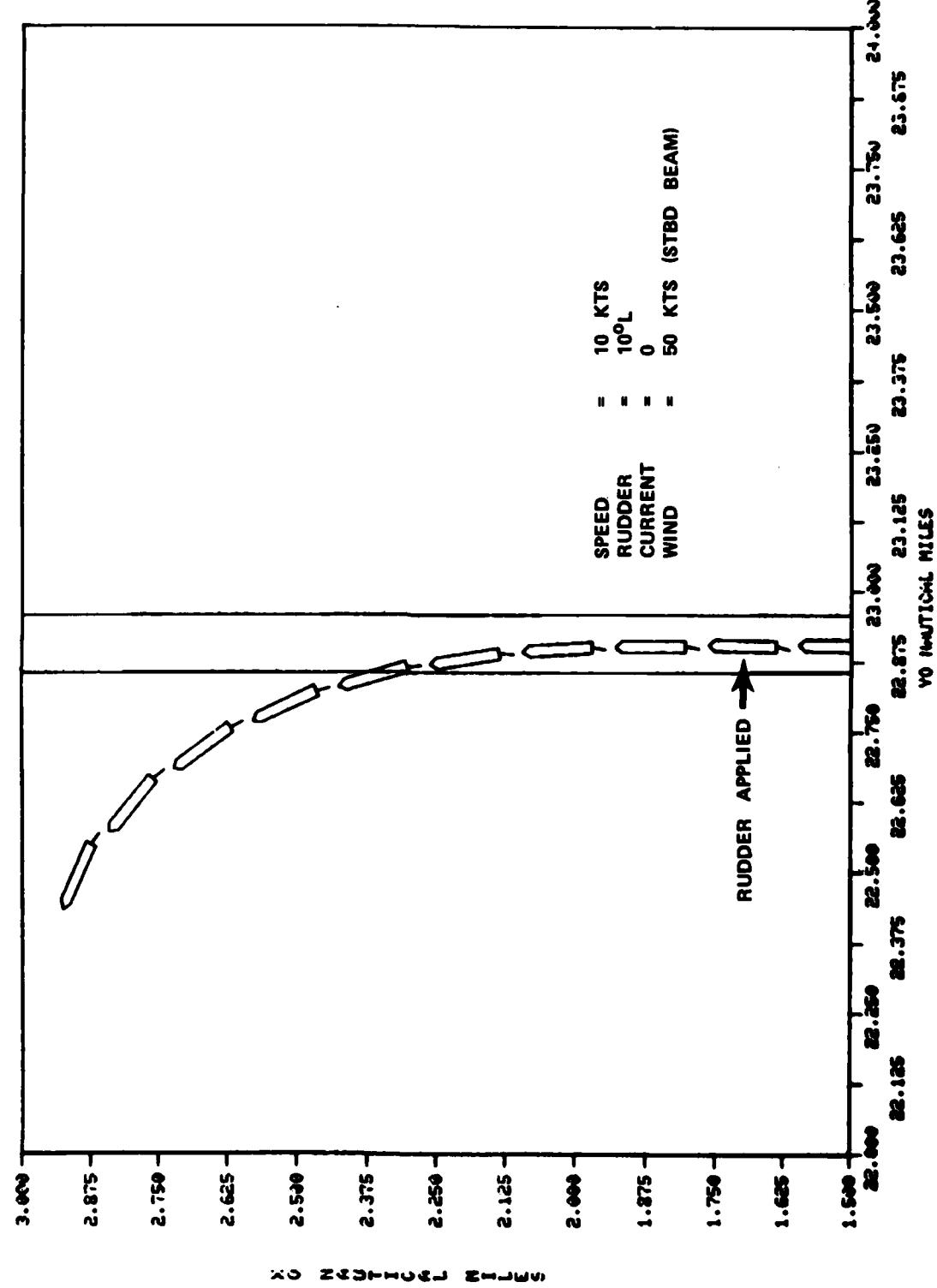
APPENDIX I

ANALYSIS TECHNIQUES

Experimental Design "A"

General. Each of the six main effects was aliased with strings of two-factor, three-factor, and higher order interactions (see Table I-1). For example, column 2 indicates that effect 2 is composed of main affect "B" plus the following interactions ACD, CE, ABDE, ABCF, DF, AEF, and BCDEF.

TABLE I-1. INDEPENDENT FACTORAL EFFECTS AND ALIASED INTERACTIONS[1]

Effects						
1	**2**	**3**	**4**	**5**	**6**	**7**
ACDEF	BCDEF	DEF	ABDEF	BDEF	ADEF	ABCDEF
BEF	AEF	ABCEF	CEF	ACEF	BCEF	EF
ABDF	DF	BCDF	ACDF	CDF	ABCDF	ADF
CF	ABCF	AF	BF	ABF	F	BCF
DL	ABDE	ACDE	BCDE	ABCDE	CDE	BDE
ABCE	CE	BE	AE	E	ABE	ACE
BCD	ACD	ABD	D	AD	BD	CD
A	B	C	ABC	BC	AC	AB

As the experimenter proceeds with the screening process, he must determine whether to neglect these interactions or investigate them. Simon makes the following generalizations concerning interactions:

- Four-factor interactions and higher are negligible for all practical purposes.

- In over 75 percent of the experiments, three-factor interaction effects can be considered negligible. However, as the number of variables studied in an experiment decrease, some three-way interactions are large enough to require further examination. Cochran and Cox suggest watching the two-factor interactions for clues that the three-factor interactions

[1] Charles Simon, **Economical Multifactor Designs for Human Factors Engineering Experiments**, Hughes Aircraft Company, 1973, p. 96.

might be important. They suggest that if the main effects and two-factor interactions of a set of factors are large, it is likely that some three-factor interactions might also be large. If the two-factor interactions are small, it is less likely (but not impossible) that the three-factor interactions are large.

● Two-factor interactions, in general, cannot a priori be assumed negligible.

For the purposes of this experiment, the main effects shall be assumed to be contributed to the experimental variables. If a given experimental variable exhibits a large proportion of the total variance, then the possible two-factor interactions are analyzed. If it is suspected that one of the aliased two-factor interactions contributed significantly to the effect, then this perspective is documented in the discussion of the effect contained in Section 2.5, Data Analysis.

Magnitude of Effects Calculations. The main effect of each variable was estimated by averaging all the performance scores obtained when the "high" condition was tested and subtracting the average of all the performance scores obtained when the "low" condition was tested. The equations for the main effects of the six experimental variables in experimental design "A" were as follows:

$$\text{Variable A} = \left(\frac{G2 + G3 + G6 + G7}{4}\right) - \left(\frac{G1 + G4 + G5 + G8}{4}\right)$$

$$\text{Variable B} = \left(\frac{G4 + G5 + G6 + G7}{4}\right) - \left(\frac{G1 + G2 + G3 + G8}{4}\right)$$

$$\text{Variable C} = \left(\frac{G2 + G5 + G6 + G8}{4}\right) - \left(\frac{G1 + G3 + G5 + G7}{4}\right)$$

$$\text{Variable D} = \left(\frac{G3 + G5 + G6 + G8}{4}\right) - \left(\frac{G1 + G2 + G4 + G7}{4}\right)$$

$$\text{Variable E} = \left(\frac{G3 + G4 + G7 + G8}{4}\right) - \left(\frac{G1 + G2 + G5 + G6}{4}\right)$$

$$\text{Variable F} = \left(\frac{G2 + G5 + G7 + G8}{4}\right) - \left(\frac{G1 + G3 + G4 + G6}{4}\right)$$

Analysis of Variance. The analysis of variance (ANOVA) was used to determine the probability that the mean of the "high" condition of the experimental variable differs from the mean of the "low" condition merely by sampling error. Variance is based upon the deviation of group means about the grand mean. This is called a between groups estimate of population variance. Another variance estimate is determined by the deviation of scores within each group about their group mean. This is known as the within groups estimate of the population variance. The null hypothesis being tested is that in the population all the group means are equal. The variance estimates are distributed as an F function with degrees of freedom equal to the degrees of freedom for numerator and denominator, respectively.

$$F = \frac{S^2 \text{ between}}{S^2 \text{ within}}$$

$$S^2 = \text{variance of sample}$$

The percentiles of the F distribution are used to determine the probability of obtaining a ratio of this size merely by sampling error. One must first select an alpha level and then determine if the values of F are greater than the values in the F table (i e., the probability of obtaining this F level by chance). If so, the null hypothesis would be rejected for the particular alpha selected.

I-2

Proportion of Variance. The percentage of variance explained by each experimental variable was calculated by taking the total sum of the squares attributed to each experimental variable and dividing it by the total sum of the squares of all the variables in the experiment. Since this design involves factors at only two levels, each source has only one degree of freedom. Therefore, the sum of squares and the variance for each effect are equal. For convenience, the results presented in Section 7 have been ranked in decreasing order based on the proportion of variance which they explain.

Mann Whitney U Test. The Mann Whitney U test may be used to test whether two independent groups have been drawn from the same population, when at least ordinal measurement has been achieved. This is one of the most powerful of the nonparametric tests, and it is the most useful alternative to the parametric t test.

To apply the U test, the observations or scores from both groups must be combined and then ranked in order of increasing size. The value of U is given by the number of times that a score in one group with n_2 cases precedes a score in the other group with n_1 cases in the ranking.

To find U, the following formula is used:

$$U = n_1 n_2 + \frac{n_1 (n_1 + 1)}{2} - R_1$$

where:

n_1 = the number of cases in the smaller of two independent groups

n_2 = the number of cases in the larger of two independent groups

R_1 = the sum of ranks assigned to the group whose sample size is n; once the U value is known, the value of z is found by using the following formula:

$$z = \frac{U - \frac{n_1 n_2}{2}}{\sqrt{\left(\frac{n_1 n_2}{N (N - 1)}\right)\left(\frac{N^3 - N}{12} - \Sigma T\right)}}$$

$$N = n_1 + n_2$$

$$T = \frac{t^3 - t}{12}$$

(where t is the number of observations tied for a given rank)

The z value as given by this formula is then used to determine the probability level through use of a reference table

Fisher Exact Probability Test. This test is an extremely useful nonparametric technique for analyzing either nominal or ordinal data when the two independent samples are small in size. It is used when the scores from two independent random samples all fall into one or the other of two mutually exclusive classes. The scores are represented by frequencies in a 2 x 2 contingency table, such as:

Variable	Improved	Remained Same	Total
1	A	B	A+B
2	C	D	C+D
Total	A+C	B+D	N

A, B, C, D = frequencies

The exact probability of observing a particular set of frequencies is given by

$$P = \frac{(A+B)!\ (C+D)!\ (A+C)!\ (B+D)!}{N!\ A!\ B!\ C!\ D!}$$

Experimental Design "B"

Homogeneity of Variance Test. The probability that the variance of experimental group 9 (reduced bridge) differed from the variance of group 3 (full bridge) merely by sampling error was determined by the homogeneity of variance test. This test assumes that the variance estimates are distributed as an "F" distribution with degrees of freedom equal to the degrees of freedom for the numerator and denominator respectively. The percentiles of the F distribution are used to determine the probability of obtaining a ratio of this size merely by sampling error. One must first select an alpha level and then determine if the values of F are greater than the values in the F table (i.e., the probability of obtaining this F level by chance). If so, the null hypothesis would be rejected for the particular alpha selected.

Difference of Means Test. The probability that the mean of experimental group 3 (full bridge) differed from the mean of experimental group 9 (reduced bridge) merely by sampling error was determined by the t-test. For small sample sizes (n < 30), it is assumed that the means of the samples are distributed as a "t" distribution with degrees of freedom equal to the combined sample size minus two. The percentiles of the "t" distribution are used to determine the probability of obtaining a ratio of this size merely by sampling error. One must first select an alpha level and then determine if the values of "t" are greater than the values in the "t" table (i.e., the probability of obtaining this "t" level by chance). If so, the null hypothesis would be rejected for the particular alpha selected.

APPENDIX J

TABULAR RESULTS OF EXPERIMENTAL DESIGNS "A" AND "B"

J.1 GENERAL

The results of the two analyses completed (one for Design "A" and one for Design "B" (see Section 2.2.3)) are organized by performance measures within the following framework:

- Simulator Design Variables (Experimental Design "A")

  - Ship Motion Performance Measures

  - Human Factors Performance Measures

- Composite Bridge Fidelity (Experimental Design "B")

  - Ship Motion Performance Measures

  - Human Factors Performance Measures

For Simulator Design Variables (Design "A") the results of each individual performance measure within a given geographic leg of the test scenario are contained on a single "Summary Data Analysis Sheet" (see Table J-1). The headings for each column are defined below:

- VARIABLE: Experimental Variables

  A — Target Maneuverability

  B — Color Visual Scene

  C — Feedback Methodology

  D — Time of Day

  E — Horizontal Field of View

  F — Instructor

- LO LEVEL: Mean value of the four groups of test subjects who were trained at the low fidelity level of the variable, as measured on the test scenario, using the stated performance measure.

- HI LEVEL: Mean value of the four groups of test subjects who were trained at the high fidelity level of the variable, as measured on the test scenario, using the stated performance measure.

- (Δ) DIFFERENCE: The effect of the variable, or the differential between the "HI" and "LO" fidelity levels.

- ANOVA: The results of the analysis of variance (ANOVA), which is the probability (p) that the mean of the "HI" condition of the experimental variable differs from the mean of the "LO" condition merely by sampling error. ($p > 0.20$ is not tabulated.)

- % VARIANCE: The percentage of variance explained by each experimental variance explained by each experimental variable. It should be noted that the variables are ranked in decreasing order of the percentage of variance explained.

For Composite Bridge Fidelity (Full Bridge versus Reduced Bridge Comparison), the results of each group's performance measured within a given geographic leg of the test scenario are contained in a single table. Four tables have been allocated for each "Summary Data Analysis Sheet." The heading for each column are defined below:

- Group 3: Full Bridge Experimental Group

- Group 9: Reduced Bridge Experimental Group

- $\bar{x}$: The mean value of the six test subjects in the specified experimental group, as measured on the test scenario, using the stated performance measure

- σ: The standard deviation of the six test subjects in the specified experimental group, as measured on the test scenario using the stated performance measure

- n: The sample size, or the number of test subjects in the specified experimental group

- F: The F-statistic for the homogeneity of variance test. The probability (p) that the variances of the two experimental groups differ merely by sampling error is given in the lower right-hand box. ($p > 0.20$ is not tabulated.)

- t: The t-statistic for the difference of means test. The probability (p) that the means of the two experimental groups differ merely by sampling error is given in the lower right-hand box. ($p > 0.20$ is not tabulated.)

J.2 SIMULATOR DESIGN VARIABLES (EXPERIMENTAL DESIGN "A")

J.2.1 SHIP MOTION PERFORMANCE MEASURES — RESULTS. Three basic types of tables are used to present the ship motion performance measure results as three distinct analysis techniques (i.e., ANOVA, Mann-Whitney U Test, Fisher Test) were used. Immediately preceding the first of the series of each of the three types of tables is an example to aid the reader in interpreting the format for that particular type of table.

J.2.1.1 ANOVA Results

Example: Mean Distance From Channel Centerline, Leg 1

1. The magnitude of effect for the high and low levels of each of the six variables indicates the following:

a. Instructor (F) — From pretest to posttest, the groups trained by Instructor "A" were 80.95 feet closer towards the channel centerline that the groups trained by Instructor "B."

b. Field of View (E) — From pretest to posttest, the groups trained with the 120-degree field of view were 59.61 feet closer towards the channel centerline than the groups trained with a 240-degree field of view.

c. Visual Scene (B) — From pretest to posttest, the groups trained with the color presentation were 36.47 feet closer towards the channel centerline than the groups trained with the black and white presentation.

d. Time of Day (D) — From pretest to posttest, the groups trained at night were 30.45 feet closer towards the channel centerline than the groups trained at day.

e. Type of Feedback (C) — From pretest to posttest, the groups trained with the augmented feedback were 23.91 feet closer towards the channel centerline than the groups trained with nonaugmented feedback.

f. Target Maneuverability (A) — From pretest to posttest, the groups trained with the independent target were 7.51 feet closer towards the channel centerline than the groups trained with the canned targets.

2. Of the above six variables, the change in the magnitude of the effect of two variables — instructor ("A," ANOVA, $p \leqslant 0.02$) and field of view (120 degrees, ANOVA, $p \leqslant 0.08$) — were significant.

3. The two variables that accounted for most of the explained variance were the instructor variable (F) which explained 11.53 percent of the variance and the field of view variable that explained 6.26 percent of the variance. The remaining four variables combined accounted for approximately 4 percent of the variance.

J.2.1.2 Mann-Whitney U Test Results

Example: Composite (Graphic), Leg 1

1. The ranking scores of the high and low levels of each of the six variables indicate the following:

a. Field of View (E) — The score for the students trained with 120-degree field of view (580.5) was greater than the score for the students trained with 240-degree field of view (500.5).

b. Time of Day (D) — The score for the students trained at day (553.5) was greater than the score for the students trained at night (527.5).

c. Instructor (F) — The score for the students trained by Instructor "A" (565.0) was greater than the score for the students trained by Instructor "B" (516.0).

d. Type of Feedback (C) — The score for the students trained with the nonaugmented feedback (585.0) was greater than the score for the students trained with the augmented feedback (496).

e. Visual Scene (B) — The score for the students trained with the color presentation (557.0) was greater than the score for the students trained with the black and white presentation (524.0).

f. Target Maneuverability (A) — The score for the students trained with the canned targets (555.5) was greater than the score for the students trained with the independent targets (525.5).

2. No significant difference (i.e., $p \leqslant 0.10$) was found to exist regarding the difference between the ranking scores for the high and low level of any of the six variables.

J.2.1.3 Fisher Test Results

Example: Pass/Fail, Leg 1

1. The comparison of the pretest results with the posttest results for each of the six variables indicates the following:

a. Time of Day (D) — Eighteen of the subjects trained with the high level (day) improved; whereas 13 of the subjects trained with the low level (night) improved.

b. Visual Scene (B) — Seventeen of the subjects trained with the low level (black and white) improved; whereas 14 of the subjects trained with the high level (color) improved.

c. Type of Feedback (C) — Sixteen of the subjects trained with the low level (nonaugmented) improved; whereas 15 of the subjects trained with the high level (augmented) improved.

d. Instructor (F) — Sixteen of the subjects trained by the low level ("B") improved; whereas 15 of the subjects trained by the high level ("A") improved.

e. Target Maneuverability (A) — Sixteen of the subjects trained with the low level (canned) improved; whereas 15 of the subjects trained with the high level (independent) improved.

f. Field of View (E) — Sixteen of the subjects trained with the high level (240 degrees) improved; whereas 15 of the subjects trained with the low level (120 degrees) improved.

2. Of the above six variables, the difference between the number of subjects who improved when trained with either the high or low level, was significant for the Time of Day variable (day; Fisher, $p \leqslant 0.03$).

J.2.2 HUMAN FACTORS PERFORMANCE MEASURES — RESULTS

Example: Rudder Order Frequency, Leg 1

1. The magnitude of effect for the high and low levels of each of the six variables indicates the following:

a. Field of View (E) — The groups trained with the 120-degree field of view increased their rudder order frequency, from pretest to posttest, more than those groups trained with a 240-degree field of view. This difference was 0.62 rudder orders per minute.

b. Visual Scene (B) — The groups trained with the black and white presentation increased their rudder order frequency, from pretest to posttest, more than those groups trained with the color presentation. This difference was 0.58 rudder orders per minute.

c. Time of Day (D) — The groups trained at night increased their rudder order frequency, from pretest to posttest, more than those groups trained at day. This difference was −0.10 rudder orders per minute.

d. Target Maneuverability (A) — The groups trained with the independently maneuverable target increased their rudder order frequency, from pretest to posttest, more than the groups trained with the canned target. This difference was 0.06 rudder orders per minute.

e. Instructor (F) — The groups trained by Instructor "A" increased their rudder order frequency, from pretest to posttest, more than the groups trained by Instructor "B." This difference was 0.05 rudder orders per minute.

f. Type of Feedback (C) — The groups trained with the augmented feedback increased their rudder order frequency, from pretest to posttest, more than the groups trained with the nonaugmented feedback. This difference was 0.03 rudder orders per minute.

2. The change in the magnitude of effect for two of the above six variables proved significant:

a. Field of View (120 degree, ANOVA, $p \leqslant 0.01$)

J-4

b. Visual Scene (black and white, ANOVA, $p \leqslant 0.01$)

3. The field of view variable (E) explained 14.54 percent of the variance; whereas the visual scene variable (B) accounted for 12.89 percent of the variance. The remaining four variables combined accounted for approximately 0.5 percent of the variance.

J.3 COMPOSITE BRIDGE FIDELITY (EXPERIMENTAL DESIGN "B")

J.3.1 SHIP MOTION PERFORMANCE MEASURES — RESULTS. The ship motion performance measures results were analyzed using the homogeneity of variance test (F-test) and the difference between the means test (t-test). An example of how to interpret the following set of tables is provided below.

Example: Distance from Channel Centerline, Leg 1

1. Group 3 had a mean training gain of -9.6 feet from channel centerline; whereas Group 9 had a mean training gain of 110.2 feet. A significant difference did not exist between these two means (t = 1.101).

2. Group 3 had a variance of 122 feet; whereas Group 9's variance was 216.4 feet. No significant difference existed between these two variances (F = 3.147; df = 5, 4).

J.3.2 HUMAN FACTORS PERFORMANCE MEASURES — RESULTS

Example: Engine Order Frequency, Leg 1

1. The full bridge group (3) decreased its engine order frequency, from pretest to posttest, by 0.141 engine orders per minute. The reduced bridge group (9) decreased their engine order frequency, from pretest to posttest, by 0.012 engine oders per minute. No significant difference was found between these two group means (t = -0.20).

2. The full bridge group (3) had a standard deviation of 0.545 engine orders per minute. The reduced bridge group (9) had a standard deviation of 0.348 engine orders per minute. No significant difference was found between these group variances (F = 2.452; df = 5, 4).

TABLE J-1. SUMMARY DATA ANALYSIS SHEET

PM: Mean Distance from Channel Centerline				Leg: 1	
	Magnitude of Effect				
Variable	**LO Level**	**HI Level**	**(Δ) Difference**	**ANOVA**	**% Variance**
(F) Instructor	72.96	153.91 x	80.95	$p < 0.02$	11.53
(E) Field of View	143.24 x	83.63	-59.61	$p < 0.08$	6.26
(B) Color Visual Scene	95.20	131.67 x	36.47		1.75
(D) Time of Day	128.66 x	98.21	-30.45		1.63
(C) Feedback	101.48	125.39 x	23.91		0.89
(A) Target Maneuverability	109.68	117.19 x	7.51		0.09

TABLE J-2. SUMMARY DATA ANALYSIS SHEET
EXPERIMENTAL DESIGN "A"

PM: Mean Distance from Channel Centerline				Leg: 2	
	Magnitude of Effect				
Variable	**LO Level**	**HI Level**	**(Δ) Difference**	**ANOVA**	**% Variance**
(F) Instructor	0.00	2.71 x	2.71		2.81
(E) Field of View	0.00	2.70 x	2.70		2.80
(A) Target Maneuverability	2.33 x	0.37	-1.96		1.47
(B) Visual Scene	2.33 x	0.37	-1.96		1.47
(C) Feedback	0.37	2.33 x	1.96		1.47
(D) Time of Day	0.37	2.33 x	1.96		1.47

TABLE J-3. SUMMARY DATA ANALYSIS SHEET

PM: Mean Distance from Channel Centerline				Leg: 3	
	Magnitude of Effect				
Variable	LO Level	HI Level	(Δ) Difference	ANOVA	% Variance
(F) Instructor	−72.91	−18.22 x	54.68	p = 0.06	7.61
(D) Time of Day	−60.76	−30.38 x	30.38	p = 0.17	3.82
(A) Target Maneuverability	−60.76	−30.38 x	30.38		3.78
(C) Feedback	−60.15	−35.84 x	24.304		1.74
(B) Color Visual Scene	−54.68	−36.45 x	18.22		0.89
(E) Field of View	−51.64	−44.35 x	7.291		0.16

TABLE J-4. SUMMARY DATA ANALYSIS SHEET

PM: Mean Distance from Channel Centerline				Leg: 4	
	Magnitude of Effect				
Variable	LO Level	HI Level	(Δ) Difference	ANOVA	% Variance
(D) Time of Day	109.13 x	86.17	−22.96		2.75
(A) Target Maneuverability	106.94 x	88.37	−18.57		1.80
(F) Instructor	89.22	106.09 x	16.87		1.49
(C) Feedback	104.54 x	90.76	−13.78		0.99
(E) Field of View	91.82	103.49 x	11.67		0.71
(B) Color Visual Scene	96.95	98.36 x	1.41		0.01

TABLE J-5. SUMMARY DATA ANALYSIS SHEET

PM: Mean Distance from Recommended Track				Leg: 1	
Variable	Magnitude of Effect			ANOVA	% Variance
	LO Level	HI Level	(Δ) Difference		
(F) Instructor	175.68	274.63 x	98.95	p ≤ 0.03	10.66
(C) Feedback	197.36	252.95 x	55.59		3.36
(E) Field of View	247.4 x	202.91	−44.49		2.16
(D) Time of Day	228.47 x	221.84	−6.63		0.04
(A) Target Maneuverability	226.32 x	223.99	−2.33		0.01
(B) Color Visual Scene	224.37	225.94 x	1.57		0.0

TABLE J-6. SUMMARY DATA ANALYSIS SHEET

PM: Mean Distance from Recommended Track				Leg: 4	
Variable	Magnitude of Effect			ANOVA	% Variance
	LO Level	HI Level	(Δ) Difference		
(F) Instructor	29.68	119.65 x	89.97	p ≤ 0.01	15.85
(C) Feedback	89.18 x	60.15	−29.03		1.65
(E) Field of View	68.68	80.65 x	11.97		0.28
(D) Time of Day	71.70	77.63 x	6.93		0.07
(B) Color Visual Scene	72.11	77.22 x	5.11		0.05
(A) Target Maneuverability	73.59	75.74 x	2.15		0.01

TABLE J-7. SUMMARY DATA ANALYSIS SHEET

PM: Closest Point of Approach (CPA*)				**Leg: 1**	

	Magnitude of Effect				
Variable	**LO Level**	**HI Level**	**(Δ) Difference**	**ANOVA**	**% Variance**
(A) Target Maneuverability	136.59 x	79.65	-59.94	p < 0.01	15.33
(D) Time of Day	94.37	121.88 x	27.51	p < 0.18	3.58
(F) Instructor	117.31 x	98.93	-18.38		1.6
(B) Color Visual Scene	113.67 x	102.57	-11.10		0.58
(E) Field of View	104.59	111.65 x	7.06		0.24
(C) Feedback	110.02 x	106.23	-3.79		0.07

*Posttest score . . . not training effectiveness

TABLE J-8. SUMMARY DATA ANALYSIS SHEET

PM: Closest Point of Approach (CPA*)				**Leg: 4**	

	Magnitude of Effect				
Variable	**LO Level**	**HI Level**	**(Δ) Difference**	**ANOVA**	**% Variance**
(D) Time of Day	130.67	176.62 x	45.95	p < 0.08	7.43
(A) Target Maneuverability	138.61	168.68 x	30.07		3.18
(C) Feedback	166.49 x	140.80	-25.69		2.32
(F) Instructor	142.64	164.65 x	20.01		1.71
(B) Color Visual Scene	147.82	159.47 x	11.65		0.48
(E) Field of View	151.36	155.93 x	4.57		0.07

*Posttest scores . . . not training effectiveness

TABLE J-9. SUMMARY DATA ANALYSIS SHEET

PM: Composite (Mean Distance from Channel Centerline and CPA)				Leg: 1	
	Magnitude of Effect				
Variable	LO Level	HI Level	(Δ) Difference	ANOVA	% Variance
(F) Instructor	190.28	252.84 x	62.56	p ≤ 0.06	7.42
(E) Field of View	247.83 x	195.28	−52.53	p ≤ 0.11	5.23
(A) Target Maneuverability	246.27 x	196.85	−49.42		4.62
(B) Color Visual Scene	208.88	234.24 x	25.36		1.2
(C) Type of Feedback	211.50	231.62 x	20.12		0.76
(D) Time of Day	223.03 x	220.08	−2.94		0.01

TABLE J-10. SUMMARY DATA ANALYSIS SHEET
EXPERIMENTAL DESIGN "A"

PM: Composite (Mean Distance from Channel Centerline and CPA)				Leg: 4	
	Magnitude of Effect				
Variable	LO Level	HI Level	(Δ) Difference	ANOVA	% Variance
(F) Instructor	172.88	282.42 x	109.54	p ≤ 0.02	11.2
(C) Feedback	260.40 x	194.90	−65.60	p ≤ 0.19	4.0
(D) Time of Day	196.88	258.42 x	61.54	p ≤ 0.16	3.53
(B) Visual Scene	214.44	240.85 x	26.41		0.65
(A) Target Maneuverability	216.93	238.38 x	21.45		0.43
(E) Field of View	218.72	236.58 x	17.86		0.3

J-10

TABLE J-11. SUMMARY DATA ANALYSIS SHEET

PM: Composite (Mean Distance from Recommended Track and CPA)				Leg: 1	
	Magnitude of Effect				
Variable	**LO Level**	**HI Level**	**(Δ) Difference**	**ANOVA**	**% Variance**
(F) Instructor	292.99	373.56 x	80.57	p ≤ 0.09	6.15
(A) Target Maneuverability	362.91 x	303.64	−59.27		3.3
(C) Feedback	307.38	359.17 x	51.79		2.54
(E) Field of View	351.99 x	314.56	−37.43		1.33
(D) Time of Day	322.83	343.72 x	28.89		0.41
(B) Visual Scene	338.04 x	328.51	−9.53		0.08

TABLE J-12. SUMMARY DATA ANALYSIS SHEET

PM: Composite (Mean Distance from Recommended Track and CPA)				Leg: 4	
	Magnitude of Effect				
Variable	**LO Level**	**HI Level**	**(Δ) Difference**	**ANOVA**	**% Variance**
(F) Instructor	231.86	270.74 x	38.88		2.59
(C) Feedback	271.04 x	231.56	−39.48		1.24
(D) Time of Day	239.80	262.79 x	22.99		0.91
(E) Field of View	243.18	259.43 x	16.25		0.45
(B) Color Visual Scene	244.77	257.83 x	13.06		0.29
(A) Target Maneuverability	245.54	257.05 x	11.51		0.23

TABLE J-13. SUMMARY DATA ANALYSIS SHEET

PM: Deviation from Desired Heading				Leg 2	
	Magnitude of Effect				
Variable	**LO Level**	**HI Level**	**(Δ) Difference**	**ANOVA**	**% Variance**
(C) Feedback	0.02	1.20 x	1.18	$p < 0.08$	6.6
(E) Field of View	1.17 x	0.05	−1.12	$p < 0.09$	5.9
(F) Instructor	1.05 x	0.16	−0.89	$p < 0.18$	3.69
(B) Color Visual Scene	0.30	0.92 x	0.62		1.76
(A) Target Maneuverability	0.66 x	0.56	−0.10		0.05
(D) Time of Day	0.66 x	0.55	−0.11		0.05

TABLE J-14. SUMMARY DATA ANALYSIS SHEET

PM: Deviation from Desired Heading				Leg: 3	
	Magnitude of Effect				
Variable	**LO Level**	**HI Level**	**(Δ) Difference**	**ANOVA**	**% Variance**
(B) Color Visual Scene	−1.79	−0.49 x	1.30	$p < 0.07$	7.26
(D) Time of Day	−1.63	−0.64 x	0.99	$p < 0.18$	4.1
(F) Instructor	−0.86 x	−1.41	−0.55		1.3
(E) Field of View	−5.97	−1.28 x	4.69		0.35
(C) Feedback	−1.24	−1.03 x	0.21		0.19
(A) Target Maneuverability	−1.11 x	−1.16	−0.05		0.01

J-12

TABLE J-15. SUMMARY DATA ANALYSIS SHEET

PM: Deviation from Desired Heading (DL40)				Leg: 2	
	Magnitude of Effect				
Variable	LO Level	HI Level	(Δ) Difference	ANOVA	% Variance
(E) Field of View	2.43 x	1.95	-0.48	p ≤ 0.003	18.3
(B) Color Visual Scene	-1.19	1.67 x	2.86	p ≤ 0.04	7.8
(A) Target Maneuverability	0.80 x	-0.32	-1.12		0.88
(F) Instructor	0.41 x	0.07	-0.34		0.1
(D) Time of Day	0.32 x	0.17	-0.15		0.02
(C) Feedback	0.21	0.41 x	0.20		0.0

TABLE J-16. SUMMARY DATA ANALYSIS SHEET

PM: Distance from Channel Centerline (DL40)				Leg: 2	
	Magnitude of Effect				
Variable	LO Level	HI Level	(Δ) Difference	ANOVA	% Variance
(E) Field of View	48.60 x	-12.15	-.0.76		6.0
(B) Color Visual Scene	42.53 x	-6.07	-48.60		4.5
(C) Feedback	-2.43	40.10 x	42.53		3.0
(F) Instructor	5.46	30.38 x	24.91		1.1
(D) Time of Day	7.29	29.77 x	22.48		0.7
(A) Target Maneuverability	18.22 x	12.15	-6.07		0.1

TABLE J-17. SUMMARY DATA ANALYSIS SHEET

PM: Distance from Channel Centerline (DL58)   Leg: 3

| Variable | Magnitude of Effect | | | ANOVA | % Variance |
	LO Level	HI Level	(Δ) Difference		
(C) Feedback	-78.98	-24.30 x	54.68	$p \leqslant 0.08$	7.3
(B) Color Visual Scene	-486.08	-36.45 x	449.62		1.7
(A) Target Maneuverability	-46.17 x	-57.11	-10.93		0.27
(F) Instructor	-55.89	-48.00 x	7.89		0.13
(E) Field of View	-53.46	-50.43 x	3.03		0.03
(D) Time of Day	-55.89	-48.00 x	7.89		0.0

TABLE J-18. MANN-WHITNEY "U" TEST

PM: Composite (Graphic)   Leg: 1

| Variable | Ranking Score | | U | Z | P |
	LO Level	HI Level			
(E) Field of View	580.5 x	500.5	224.5	-0.88	0.19
(D) Time of Day	527.5	553.5 x	300.5	0.8	0.21
(F) Instructor	516.0	565.0 x	289.0	0.54	0.29
(C) Feedback	585.0 x	496.0	243.0	-0.46	0.32
(B) Visual Scene	524.0	557.0 x	281.0	0.36	0.36
(A) Target Maneuverability	555.5 x	525.5	249.5	-0.33	0.37

TABLE J-19. MANN-WHITNEY "U" TEST

PM: Composite (Graphic)   Leg: 2

| Variable | Ranking Score | | U | Z | P |
	LO Level	HI Level			
(E) Field of View	574.0 x	461.0	185.0	−1.55	0.06
(A) Target Maneuverability	533.0 x	502.0	226.0	−0.62	0.27
(I) Instructor	529.0 x	506.0	230.0	−0.53	0.3
(D) Time of Day	540.0 x	495.0	242.0	−0.25	0.4
(B) Visual Scene	544.0 x	491.0	260.0	−0.18	0.43
(C) Feedback	525.0 x	510.0	257.0	−0.09	0.46

TABLE J-20. MANN-WHITNEY "U" TEST

PM: Composite (Graphic)   Leg: 3

| Variable | Ranking Score | | U | Z | P |
	LO Level	HI Level			
(A) Target Maneuverability	443.5 x	297.5	107.5	−2.17	0.01
(F) Instructor	280.0	461.0 x	208.0	0.96	0.16
(C) Feedback	401.5 x	339.5	149.5	−0.92	0.17
(E) Field of View	300.0	441.0 x	188.0	0.36	0.36
(B) Visual Scene	382.5 x	358.5	168.5	−0.36	0.36
(D) Time of Day	369.5	371.5 x	181.5	0.03	0.48

TABLE J-21. MANN-WHITNEY "U" TEST

PM: Composite (Graphic)			Leg: 4		
	Ranking Score				
Variable	LO Level	HI Level	U	Z	P
(E) Field of View	597.0 x	531.0	231.0	−0.96	0.17
(F) Instructor	519.0	609.0 x	309.0	0.7	0.24
(D) Time of Day	606.5 x	522.5	245.0	−0.65	0.26
(C) Feedback	599.5 x	528.5	252.0	−0.5	0.31
(B) Visual Scene	594.5 x	533.5	257.0	−0.39	0.35
(A) Target Maneuverability	564.5 x	563.5	287.0	0.24	0.4

TABLE J-22. FISHER TEST

PM: Pass/Fail				Leg: 1	
	Low Level		HI Level		
Variable	Improved	Other	Improved	Other	Fisher P
(D) Time of Day	13	11	18	4	0.03
(B) Visual Scene	17	6	14	9	0.16
(C) Feedback	16	8	15	7	0.25
(F) Instructor	16	7	15	8	0.47
(A) Target Maneuverability	16	7	15	8	0.47
(E) Field of View	15	8	16	7	0.47

TABLE J-23. FISHER TEST

Variable	PM: Pass/Fail		Leg: 2		
	Low Level		HI Level		
	Improved	Other	Improved	Other	Fisher P
(E) Field of View	7	15	1	23	0.02
(D) Time of Day	2	22	6	16	0.08
(F) Instructor	6	17	2	21	0.10
(B) Visual Scene	3	21	5	17	0.20
(C) Feedback	5	18	3	20	0.23
(A) Target Maneuverability	4	19	4	19	0.30

TABLE J-24. FISHER TEST

Variable	PM: Pass/Fail		Leg: 3		
	Low Level		HI Level		
	Improved	Other	Improved	Other	Fisher P
(E) Field of View	2	19	0	24	0.21
(C) Feedback	0	23	2	20	0.23
(A) Target Maneuverability	0	23	2	20	0.23
(F) Instructor	1	21	1	22	0.51
(B) Visual Scene	1	23	1	20	0.51
(D) Time of Day	1	23	1	20	0.51

TABLE J-25. FISHER TEST

| | Low Level | | HI Level | | |
Variable	Improved	Other	Improved	Other	Fisher P
(A) Target Maneuverability	16	8	12	11	0.02
(D) Time of Day	16	8	12	11	0.02
(F) Instructor	14	9	14	10	0.22
(C) Feedback	14	10	14	9	0.22
(B) Visual Scene	14	10	14	10	0.22
(E) Field of View	16	7	12	12	0.23

PM: Pass/Fail   Leg: 4

TABLE J-26. SUMMARY DATA ANALYSIS SHEET

| | Magnitude of Effect | | | | |
Variable	LO Level	HI Level	(Δ) Difference	ANOVA	% Variance
(E) Field of View	1.16	0.54	−0.62	$p \leqslant 0.01$	14.54
(B) Visual Scene	1.14	0.56	−0.58	$p \leqslant 0.01$	12.89
(D) Time of Day	0.89	0.79	−0.10		0.37
(A) Target Maneuverability	0.82	0.88	0.06		0.12
(F) Instructor	0.82	0.87	0.05		0.10
(C) Feedback	0.83	0.86	0.03		0.03

PM: Rudder Order Frequency   Leg: 1

TABLE J-27. SUMMARY DATA ANALYSIS SHEET

PM: Rudder Order Frequency				Leg: 3	

Variable	Magnitude of Effect			ANOVA	% Variance
	LO Level	HI Level	(Δ) Difference		
(B) Visual Scene	0.36	0.00	−0.36	p ≤ 0.09	6.51
(C) Feedback	0.09	0.27	0.18		1.66
(D) Time of Day	0.27	0.09	−0.18		1.59
(F) Instructor	0.11	0.25	0.14		0.95
(A) Target Maneuverability	0.22	0.13	−0.09		0.39
(E) Field of View	0.15	0.21	0.06		0.17

TABLE J-28. SUMMARY DATA ANALYSIS SHEET

PM: Rudder Order Frequency				Leg: 4	

Variable	Magnitude of Effect			ANOVA	% Variance
	LO Level	HI Level	(Δ) Difference		
(E) Field of View	0.65	0.12	−0.53	p ≤ 0.01	14.45
(B) Visual Scene	0.56	1.25	0.65	p ≤ 0.07	6.24
(F) Instructor	0.23	0.53	0.30	p ≤ 0.12	4.59
(D) Time of Day	0.48	0.29	−0.19		1.79
(A) Target Maneuverability	0.37	0.40	0.03		0.06
(C) Feedback	0.39	0.38	−0.01		0

TABLE J-29. SUMMARY DATA ANALYSIS SHEET

PM: Engine Order Frequency   Leg: 1

Variable	Magnitude of Effect			ANOVA	% Variance
	LO Level	HI Level	(Δ) Difference		
(B) Visual Scene	0.09	−0.11	−0.20	p ≤ 0.17	4.15
(D) Time of Day	0.08	−0.10	−0.18	p ≤ 0.22	3.30
(E) Field of View	0.08	−0.10	−0.18	p ≤ 0.23	3.20
(A) Target Maneuverability	0.06	−0.08	−0.14		1.78
(C) Feedback	0.00	−0.02	−0.02		0.05
(F) Instructor	0.00	−0.02	−0.02		0.02

TABLE J-30. SUMMARY DATA ANALYSIS SHEET

PM: Engine Order Frequency   Leg: 4

Variable	Magnitude of Effect			ANOVA	% Variance
	LO Level	HI Level	(Δ) Difference		
(D) Time of Day	−0.25	−0.04	0.21	p ≤ 0.22	3.4
(C) Feedback	−0.05	−0.24	−0.19	p ≤ 0.28	2.63
(A) Target Maneuverability	−0.06	−0.23	−0.17		2.03
(B) Visual Scene	−0.09	−0.20	−0.11		1.01
(F) Instructor	−0.19	−0.11	0.08		0.49
(E) Field of View	−0.12	−0.17	−0.05		0.14

J-20

TABLE J-31. SUMMARY DATA ANALYSIS SHEET

PM: Bow Thruster Frequency   Leg: 2

| Variable | Magnitude of Effect | | | ANOVA | % Variance |
	LO Level	HI Level	(Δ) Difference		
(A) Target Maneuverability	0.00	0.02	0.02	p ≤ 0.01	14.12
(F) Instructor	−0.01	0.03	0.04	p ≤ 0.22	2.95
(D) Time of Day	−0.01	0.04	0.05	p ≤ 0.25	2.58
(B) Visual Scene	0.04	−0.09	−0.13		1.62
(C) Feedback	−0.04	0.07	0.11		0.15
(E) Field of View	−0.04	0.06	0.10		0.02

TABLE J-32. SUMMARY DATA ANALYSIS SHEET

PM: Bow Thruster Frequency   Leg: 3

| Variable | Magnitude of Effect | | | ANOVA | % Variance |
	LO Level	HI Level	(Δ) Difference		
(E) Field of View	−0.06	0.11	0.17	p ≤ 0.01	14.64
(F) Instructor	0.07	−0.02	−0.09	p ≤ 0.14	4.41
(B) Visual Scene	0.06	0.00	−0.06		1.76
(C) Feedback	0.01	0.05	0.04		0.87
(A) Target Maneuverability	0.04	0.02	−0.02		0.23
(D) Time of Day	0.03	0.02	−0.01		0.13

TABLE J-33. SUMMARY DATA ANALYSIS SHEET

PM: Reaction Time   Leg: 2

Variable	Magnitude of Effect			ANOVA	% Variance
	LO Level	HI Level	(Δ) Difference		
(A) Target Maneuverability	−0.37	−5.91	−5.54		2.92
(B) Visual Scene	−1.54	−4.75	−3.21		0.98
(F) Instructor	−2.33	−3.95	−1.62		0.25
(D) Time of Day	−3.83	−2.45	+1.38		0.18
(C) Feedback	−3.66	−2.62	+1.04		0.10
(E) Field of View	−3.66	−2.62	+1.04		0.10

TABLE J-34. SUMMARY DATA ANALYSIS SHEET

PM: Distance from Channel Centerline				Leg: 1	
Group	x̄	σ	n	F	t
3	-9.6	122.0	6	3.147	1.101
9	+110.2	216.4	5		

PM: Distance from Channel Centerline				Leg: 2	
Group	x̄	σ	n	F	t
3	-13.7	47.7	6	3.7	0.358
9	-7.2	24.8	5		

PM: Distance from Channel Centerline				Leg: 3	
Group	x̄	σ	n	F	t
3	-115.7	155.9	6	1.5	1.339
9	-2.55	124.3	5		

PM: Distance from Channel Centerline				Leg: 4	
Group	x̄	σ	n	F	t
3	93.1	73.5	6	2.4439	0.748
9	120.5	47	5		

J-23

TABLE J-35. SUMMARY DATA ANALYSIS SHEET

PM: Distance from Recommended Track				Leg: 1	
Group	$\bar{x}$	σ	n	F	t
3	194.1	121.9	6	4.81	0.16
9	172.5	267.3	5		

PM: Distance from Recommended Track				Leg: 4	
Group	$\bar{x}$	σ	n	F	t
3	66.2	119.7	6	1.704	0.2116
9	79.7	91.7	5		

PM: Deviation from Desired Heading				Leg: 2	
Group	$\bar{x}$	σ	n	F	t
3	-0.78	2.65	6	1.2	0.398
9	-1.45	2.91	5		

PM: Deviation from Desired Heading				Leg: 3	
Group	$\bar{x}$	σ	n	F	t
3	-0.22	3.29	6	1.1327	1.98
9	-4.3	3.5	5		$\alpha = 0.05$

J-24

TABLE J-36. SUMMARY DATA ANALYSIS SHEET

PM: Distance from Channel Centerline (DL40)					Leg: 2
Group	$\bar{x}$	σ	n	F	t
3	−54.88 feet	163.8	6	15.7	0.2669
9	−73.52 feet	41.3	4	$\alpha = 0.05$	

PM: Heading Difference (DL40)					Leg: 2
Group	$\bar{x}$	σ	n	F	t
3	4.8	3.806	6	1.473	2.129
9	0.1	3.136	4		$\alpha = 0.05$

TABLE J-37. SUMMARY DATA ANALYSIS SHEET

PM: Distance from Channel Centerline (DL58)				Leg: 3	
Group	$\bar{x}$	σ	n	F	t
3	−99.5 feet	1054.2	5	79.98	1.898
9	−92.9 feet	117.9	4	$\alpha = 0.01$	

PM: CPA* (Tug and Tow)				Leg: 1	
Group	$\bar{x}$	σ	n	F	t
3	119.4	94.05	6	1.6898	0.4783
9	84.88	122.26	5		

PM: CPA* (Tug and Tow)				Leg: 4	
Group	$\bar{x}$	σ	n	F	t
3	234.33	100.15	6	2.734	1.566
9	157.52	60.57	5		$\alpha = 0.10$

*Posttest score . . . not training effectiveness

TABLE J-38. SUMMARY DATA ANALYSIS SHEET

PM: Engine Order Frequency				Leg: 1	
Group	$\bar{x}$	σ	n	F	t
3	−0.141	0.545	6	2.452	−0.20
9	−0.012	0.348	5		

PM: Engine Order Frequency				Leg: 4	
Group	$\bar{x}$	σ	n	F	t
3	0.028	0.439	6	1.520	0.34
9	−0.164	0.356	5		

PM: Rudder Order Frequency				Leg: 1	
Group	$\bar{x}$	σ	n	F	t
3	1.030	0.445	6	1.612	0.28
9	0.826	0.565	5		

PM: Rudder Order Frequency				Leg: 3	
Group	$\bar{x}$	σ	n	F	t
3	0.096	0.603	6	2.963	0.20
9	−0.138	1.038	5		

J-27

TABLE J-39. SUMMARY DATA ANALYSIS SHEET

PM: Rudder Order Frequency				Leg: 4	
Group	x̄	σ	n	F	t
3	0.362	0.785	6	7.293	0.047
9	0.256	2.120	5	α = 0.05	

PM: Bow Thruster Frequency				Leg: 2	
Group	x̄	σ	n	F	t
3	0.070	0.120	6	2.141	0.74
9	−0.038	0.082	5		

PM: Bow Thruster Frequency				Leg: 3	
Group	x̄	σ	n	F	t
3	0.103	0.237	6	1.157	0.61
9	−0.110	0.255	5		

PM: Reaction Time				Leg: 2	
Group	x̄	σ	n	F	t
3	6.83	20.64	6	3.93	−0.74
9	−27.0	40.94	5	α = 0.05	

APPENDIX K

ANALYSIS OF DEBRIEFING QUESTIONNAIRE

INTRODUCTION

This appendix is comprised of three sections. Section 1 presents a sample of the debriefing questionnaire used during the individual interviews with each of the 52 training program participants. The questionnaire consists of 21 questions which require the trainees to rate their answers on a 1 through 10 scale, with 1 being the lowest rating and 10 being the highest. Section 2 specifies in tabular form, the statistical results obtained from the analysis of the answers given to the 21 questions posed. Two tables are presented. Table 1 lists the mean and standard deviation for each question based upon a compilation of all 52 trainees' responses. Table 2 lists the mean and standard deviation of each of the nine group responses to Question 18, which addresses the degree to which a specific lower fidelity characteristic (which is identified in the table in reference to each group) interfered with training. Section 3 presents, in narrative form, the results of the analysis and a discussion of the results.

SECTION 1

DEBRIEFING QUESTIONNAIRE

GROUP   DATE

NAME

SUBJECT #

1) CIRCLE A WEIGHT ASSOCIATED WITH YOUR PERFORMANCE ON THE PRETEST.
    1   2   3   4   5   6   7   8   9   10
    COMMENTS:

2) CIRCLE A WEIGHT ASSOCIATED WITH YOUR PERFORMANCE ON THE POSTTEST.
    1   2   3   4   5   6   7   8   9   10
    COMMENTS:

3) CIRCLE A WEIGHT ASSOCIATED WITH THE EFFECTIVENESS OF THE TRAINING PROGRAM.
    1   2   3   4   5   6   7   8   9   10
    COMMENTS:

4) CIRCLE A WEIGHT WHICH IS ASSOCIATED WITH YOUR RELIANCE ON INFORMATION OBTAINED FROM ELECTRONIC AIDS (I.E., RADAR, RATE OF TURN INDICATOR, GYRO COMPASS) DURING THE TRAINING PROGRAM.
    1   2   3   4   5   6   7   8   9   10
    COMMENTS:

5) CIRCLE A WEIGHT WHICH BEST DESCRIBES YOUR ASSESSMENT OF THE HANDLING CHARACTERISTICS OF THE SHIP.
    1   2   3   4   5   6   7   8   9   10
    COMMENTS:

6) CIRCLE A WEIGHT WHICH IS ASSOCIATED WITH HOW ADEQUATE YOU FEEL THE AMOUNT OF TIME SPENT ON TRAINING TURNS WAS.
    1   2   3   4   5   6   7   8   9   10
    COMMENTS:

7) CIRCLE A WEIGHT WHICH IS ASSOCIATED WITH HOW ADEQUATE YOU FEEL THE AMOUNT OF TIME SPENT ON TRAINING RUDDER FAILURES WAS.
    1   2   3   4   5   6   7   8   9   10
    COMMENTS:

8) CIRCLE A WEIGHT WHICH IS ASSOCIATED WITH HOW ADEQUATE YOU FEEL THE AMOUNT OF TIME SPENT ON TRAINING WIND/CURRENT EFFECTS WAS.
    1   2   3   4   5   6   7   8   9   10
    COMMENTS:

9) CIRCLE A WEIGHT WHICH IS ASSOCIATED WITH HOW ADEQUATE YOU FEEL THE AMOUNT OF TIME SPENT ON TRAINING AID TO NAVIGATION DISCREPANCIES WAS.

1 2 3 4 5 6 7 8 9 10

COMMENTS:

10) CIRCLE A WEIGHT WHICH IS ASSOCIATED WITH HOW ADEQUATE YOU FEEL THE AMOUNT OF TIME SPENT ON TRAINING POWER FAILURES WAS.

1 2 3 4 5 6 7 8 9 10

COMMENTS:

11) CIRCLE A WEIGHT WHICH BEST DESCRIBES THE EFFECTIVENESS OF THE CLASSROOM TRAINING.

1 2 3 4 5 6 7 8 9 10

COMMENTS:

12) CIRCLE A WEIGHT WHICH BEST DESCRIBES THE EFFECTIVENESS OF THE CLASSROOM AIDS (E.G., TRANSPARENCIES, SLIDES).

1 2 3 4 5 6 7 8 9 10

COMMENTS:

13) CIRCLE A WEIGHT WHICH BEST DESCRIBES THE EFFECTIVENESS OF THE SHIP PLOTS AS A FEEDBACK DEVICE.

1 2 3 4 5 6 7 8 9 10

COMMENTS:

14) CIRCLE A WEIGHT WHICH BEST DESCRIBES THE EFFECTIVENESS OF THE TV/SITUATION DISPLAY AS A FEEDBACK DEVICE.

1 2 3 4 5 6 7 8 9 10

COMMENTS:

15) CIRCLE A WEIGHT WHICH BEST DESCRIBES HOW MUCH SHIPHANDLING YOU LEARNED IN THE TRAINING PROGRAM.

1 2 3 4 5 6 7 8 9 10

COMMENTS:

16) CIRCLE A WEIGHT WHICH BEST DESCRIBES THE TRAINING EFFECTIVENESS OF THE TRAINING THAT YOU RECEIVED BY OBSERVING THE OTHER PARTICIPANTS CONNING THE SHIP.

1 2 3 4 5 6 7 8 9 10

COMMENTS:

17) CIRCLE A WEIGHT WHICH BEST DESCRIBES THE TRAINING EFFECTIVENESS OF THE TRAINING THAT YOU RECEIVED BY CONNING THE SHIP YOURSELF.

1 2 3 4 5 6 7 8 9 10

COMMENTS:

18) CIRCLE A WEIGHT WHICH IS ASSOCIATED WITH HOW MUCH YOU FEEL THE BLACK AND WHITE IMAGE INTERFERED WITH YOUR TRAINING.

1 2 3 4 5 6 7 8 9 10

COMMENTS:

19) CIRCLE A WEIGHT WHICH BEST DESCRIBES THE AMOUNT OF AWARENESS THAT YOU HAVE GAINED IN UNDERSTANDING WHAT A MASTER OR PILOT IS FACED WITH WHEN HE MANEUVERS A VESSEL.

    1   2   3   4   5   6   7   8   9   10

    COMMENTS:

20) CIRCLE A WEIGHT THAT IS ASSOCIATED WITH THE AMOUNT OF CONFIDENCE THAT YOU HAVE GAINED IN DEALING WITH EMERGENCY SITUATIONS IN THE REAL WORLD.

    1   2   3   4   5   6   7   8   9   10

    COMMENTS:

21) WAS THE EQUIPMENT USED SIMILAR TO YOUR RECENT EXPERIENCE, FOR INSTANCE:

    Radars

    Wind Indicators

    Fathometers

    Radios

    Engine room telegraphs

    General layout

    Bow thruster

SECTION 2

STATISTICAL RESULTS

TABLE 1. STATISTICAL RESULTS OF ALL SUBJECTS' RESPONSES TO EACH QUESTIONNAIRE ITEM

Question Number	Mean	Standard Deviation
1	3.28	1.57
2	8.01	1.42
3	8.84	1.25
4	7.51	1.71
5	7.17	2.47
6	7.98	2.05
7	7.46	2.16
8	8.34	1.70
9	6.53	2.50
10	7.78	2.04
11	7.26	2.89
12	8.01	2.19
13	7.48	2.59
14	8.56	2.55
15	8.19	1.84
16	8.67	1.53
17	9.11	1.35
18	4.71	3.10
19	8.69	1.99
20	8.45	1.76
21	Calculation not appropriate	1.76

TABLE 2. STATISTICAL RESULTS FOR EACH GROUP'S RESPONSES TO QUESTION 18, THE DEGREE TO WHICH THE LOWER FIDELITY CHARACTERISTICS OF SIMULATION INTERFERED WITH TRAINING

Group	Low Fidelity Characteristic	Mean	Standard Deviation
A	Black and White Visual Scene	3.3	1.63
B	120° Horizontal Field of View	3.5	3.39
C	Night	6.5	3.61
D	Black and White/Night	4.16	3.86
E	120° Horizontal Field of View	7.0	1.87
F	Night	4.66	3.66
G	Black and White/Night	3.83	3.37
H	Black and White Visual Scene	4.16	2.99
I	Black and White Visual Scene	5.0	2.44

SECTION 3

DEBRIEFING QUESTIONNAIRE RESULTS AND DISCUSSION

Each of the 52 trainees who took part in the training program over the 9-week period, individually participated in a debriefing session conducted upon completion of his participation in the training program. The debriefing session took the form of a structured interview which was conducted to (1) evaluate the training they received in terms of its overall effectiveness, (2) obtain feedback regarding the contents of the program, (3) determine the value of participation in such a program, and (4) to generate additional data for analysis regarding the experimental objectives. During the interview, 21 questions (see Section 1, Sample Questionnaire) were posed, requiring each trainee to weight their answers on a scale of 1 through 10, with 1 being the lowest rating and 10 being the highest rating. Comments regarding the questions were also elicited. An analysis yielding a mean and a standard deviation of the weights assigned to each question was completed (see Section 1, Tables 1 and 2). The following presents both the results of this analysis and a discussion of the comments supplied for each question.

PRETEST PERFORMANCE (Question 1)

Trainees rated their pretest performance at a mean of 3.28 (1.57), attributing their relatively poor results to a lack of familiarity with the CAORF bridge layout, equipment, and ownship handling characteristics. Nervousness was cited by some as a contributing factor to their initial lack of success, while others pointed to unreal environmental effects (bank suction — leg 4, current — leg 1). Self-critique was not evident in trainee comments on their pretest evaluation, although it may be inferred in subsequent commentary regarding program element training effectiveness.

POSTTEST PERFORMANCE (Question 2) AND TRAINING PROGRAM EFFECTIVENESS (Question 3)

Posttest performance was rated by trainees as a whole at a mean of 8.01 (1.42), showing a mean subjective training gain of 4.73 (1.91) on the 10 point scale. This substantial improvement correlates well with the perceived effectiveness of the training program expressed in response to several other questions. The program was given an overall effectiveness rating of 8.84 (1.25). Most frequently cited as reasons for better performance in the posttest were a feeling of confidence and a greater familiarity with ownship's handling characteristics and environmental parameters.

RELIANCE ON INFORMATION OBTAINED FROM ELECTRONIC AIDS (Question 4)

Although the mean response of 7.51 (1.71) to the question dealing with reliance upon electronic aids is not very informative, commentary demonstrated a very strong reliance upon, and endorsement of the rate of turn indicator. Only 1 had previously sailed with such a device, but it was overwhelmingly hailed as a major aid to their performance in turns. More than one trainee suggested that its carriage at-sea be a legal requirement. Radar, to the contrary, was reported to have been employed very little as the scenarios were not of reduced visibility.

OWNSHIP HANDLING CHARACTERISTICS (Question 5)

The rating response to Question 5, which asks for an assessment of the handling characteristics of ownship, is difficult to interpret as it was not clear to trainees whether they were being asked to evaluate the fidelity of ownship's hydrodynamics or to provide a relative assessment of the maneuverability of the vessel in comparison to other ship types. A high measure of dispersion, SD = 2.47, demonstrates this ambiguity. All agreed that the vessel was sluggish and unresponsive and most imagined that this characteristic was probably realistic. Very few reported any real-world experience with a similar sized tanker (or tankers in general), but the two that did, felt that ownship had been modeled too sluggish.

TIME SPENT ON TRAINING TURNS, RUDDER FAILURES, WIND/CURRENT EFFECTS, AIDS TO NAVIGATION DISCREPANCIES, AND POWER FAILURES (Questions 6-10)

The ratings for Questions 6-10 are also somewhat difficult to interpret as the commentary demonstrates that many trainees rated the importance and not adequacy of the several training issues or topics addressed. Unfortunately, the two interpretations would lead to complimentary rather than parallel ratings. However, it is clear from the commentary that training in making turns and dealing with wind and especially current, was felt to be very important. In contrast, training relating to aids to navigation discrepancies was generally felt to be of a low order of priority, this in spite of the fact that Question 9, received the lowest mean rating of 6.58 (2.50) of the "training adequacy" queries (meaning least adequate training). The ambiguity of these questions may also be indicated by the relatively high magnitudes of measures of dispersion.

CLASSROOM TRAINING EFFECTIVENESS (Question 11) AND CLASSROOM AID EFFECTIVENESS (Question 12)

Ratings of effectiveness of classroom training resulted in an overall mean of 7.26 (2.89) which showed a high measure of dispersion due to the fact that some trainees evaluated the effectiveness of the program's classroom training against an unstated model of what it might have been (the ideal classroom training program), while others compared the effectiveness of this classroom training experience to the effectiveness of training on the bridge. Trainees were much more satisfied with instruction given during simulation than in the classroom. Other than for the first day's classroom presentation, they found it repetitive and not sufficiently generalizing or theoretical. The training was so specific to the simulation that they would have preferred to have received it on the bridge in the same vein (specifically), several trainees suggested that the classroom sessions were non-productive in that they prepared the students (put them on their guard) for specific failures to follow. They would have preferred not to have classroom discussion of emergency procedures followed immediately by simulations of those failures, feeling that an element of surprise would have improved training.

In general, the trainees opinion of the classroom training was, as one put it, "eliminate it or drastically expand it." As to classroom lecture content, trainees desired to have more information on ownship's handling characteristics.

EFFECTIVENESS OF SHIP PLOTS AS A FEEDBACK DEVICE (Question 13) AND EFFECTIVENESS OF THE TV/SITUATION DISPLAY AS A FEEDBACK DEVICE (Question 14)

Those trainees which did receive track plots of their training runs did value them but felt they were of value only if they could review them immediately after the run. The TV display was rated even more highly as a feedback device, although the opinion was widely shared that the trainee at the conn should not be allowed to view it, as such a device (display) is not available to real-world shiphandlers. It was very highly rated, however, as an aid to observers. One improvement asked for is a history of positions (track) throughout the run rather than a single moving target.

SHIPHANDLING TRAINING GAIN (Question 15)

Trainees evaluate their own shiphandling training gain at a mean of 8.19 (1.84) which corresponds with their estimates of personal pretest to posttest training gain and evaluation of the overall program training effectiveness. They point out that they came into the program with very little or no shiphandling experience in restricted waterways and with little recent open-water experience due to their non-watchstanding positions as chief mates.

TRAINING EFFECTIVENESS OF TRAINING RECEIVED BY OBSERVING THE OTHER PARTICIPANTS CONNING THE SHIP (Question 16) AND THE EFFECTIVENESS OF THE TRAINING RECEIVED BY CONNING THE VESSEL THEMSELVES (Question 17)

The training effectiveness of the observing element of the program was rated highly at a mean of 8.67 (1.53). It was noted by several trainees that they were able to learn a great deal while in observer status, in that they were not under pressure and could calmly evaluate the performance of others and learn from their mistakes. Pressure notwithstanding, however,

trainees rated the training effectiveness of personal "hands-on" control of ownship at an even higher mean rating of 9.11 (1.35). They felt that no other element of the training could compare with "doing it" oneself.

INTERFERENCE WITH TRAINING DUE TO BLACK AND WHITE IMAGE, NIGHT SCENE, REDUCED BRIDGE, AND/OR 120° HORIZONTAL FIELD OF VIEW (Question 18)

In comparison to responses for other questions, there appears to be little unanimity on the degree to which "lower fidelity" characteristics of the simulation interfered with training. Black and white simulation was generally not felt to be a serious drawback, one student comparing the effect to pre-dawn overcast conditions in the real world. It was expressed that buoys were more difficult to distinguish and buoy lights and sidelights more difficult to identify under black and white conditions. "Night" not really being a lower level of fidelity was not considered to be detrimental to training as trainees pointed out that night conditions exist also in the real world. It was felt, however, that the night and black and white combined conditions were not as stimulating and interesting as their counterparts and that the lack of visual stimulation contributed to the tedium of repetitive scenarios. The reduced bridge was similarly evaluated as interfering somewhat with training effectiveness, which is indicated by the 5.0 mean rating, but with significant difference in opinion (SD = 2.44). Especially annoying were the "blind spots" occasioned by the frames or posts which separated the several TV screens.

The two groups which trained under reduced horizontal field of view (120°) had different opinions as to its limitations as indicated by the means of Group B — 3.5 (3.39) versus Group E — 7.0 (1.87). There was general agreement that the expanded view was of assistance in making the turns and a few commented that they missed the opportunity to observe aids as they pass abeam.

AWARENESS OF THE POSITION AND RESPONSIBILITIES OF A MASTER OR PILOT (Question 19)

As to trainee perception of the amount of awareness that they gained in understanding what a pilot or master is faced with in maneuvering, many responded that they were previously aware of the demands of the task, while others reported an expanded awareness as indicated by the mean of 8.69 (1.99).

CONFIDENCE GAINED IN HANDLING EMERGENCY SITUATIONS (Question 20)

Significant improvement in self-confidence in emergency shiphandling was perceived by almost all trainees as indicated by the high rating of 8.45 (1.76).

EQUIPMENT FAMILIARITY (Question 21)

In general, the trainees were not familiar with either the layout or equipment of the CAORF bridge. In comparison to the ships in which they have sailed, they found the CAORF layout to be quite unusual with the consoles forward against the windows. The radar/CAS was much more sophisticated than models they had sailed with and the wind indicator was foreign to many. The digital fathometer was new to them as was the bow thruster and bridge control console. The radios were the only piece of gear familiar to most.

APPENDIX L

SUMMARY OF INSTRUCTOR CHARACTERISTICS

A brief search of the literature (Athay, 1974; Badura, 1975; Blimline, 1975; Garrett, 1978; Hamachek, 1972; Highet, 1950; Rose, 1961; Cronbach, 1963) for information about instructor characteristics (i.e., the attributes possessed by a "good"/ "effective" instructor) yielded the following general list of attributes based on the findings of several experimental studies:

1. Competence in subject matter being taught

2. Enjoyment of what is taught

3. Desire to teach

4. Has wide and lively intellectual interests so as to make the subject relevant and to make work more interesting for the students

5. Mastery of the techniques of instruction

   ● Uses a conventional manner in teaching — informal, easy style

   ● Speaks clearly

   ● Organizes instruction according to the learning capacities of the students

   ● Knows what his students are ready for, knows the language they understand best, and knows how fast new material can be learned

   ● Emphasizes student-led discussions and an open format rather than the traditional approach

   ● Repeats an : emphasizes key material in such a way that it stands the best chance of being remembered

   ● Conducts demonstrations skillfully; uses examples in instruction

   ● Is skilled in asking questions (as opposed to seeing oneself as a kind of answering service)

   ● The habit of evaluation — administers practice periods and tests in such a way as to promote and develop desirable skills and attitudes

   ● Resourceful and creative

     − Has a willingness to be flexible, to be direct or indirect as the situation demands

     − Does not pursue a single behavioral-instructional path to the exclusion of other possibilities

- Uses methods which provide for adaptation to individual differences, encourages student initiative, and stimulates individual and group participation

- Possesses the ability to develop good personal relationships

  - Empathetic

  - Genuine

  - Good sense of humor

  - Warm

- "Warm" teachers refer to those individuals who possess any or all of the following qualities:

  - Spontaneous expression of feeling — The teacher colors classroom relationships with a continual expression of his own enthusiasm and liking for his students

  - Supports and encourages — Reinforcement is noncontingent; the teacher approves of the pupil as a person, whatever he does, persuades him that he can reach his goal, and helps him over obstacles

  - Contingent social reinforcement — The teacher gives plentiful praise but only when he judges the pupil's actions to be meritorious; approval is given when it is earned, not otherwise

  - Tact and considerateness — Criticism or rejection of a pupil's proposal is presented in such a way that the pupil does not feel blamed or inferior

  - Acceptance of pupil's feeling — The teacher encourages the pupil to express his interests, fears, etc., and takes them seriously

APPENDIX M

TRAINING ASSISTANCE TECHNOLOGY

The U.S. Navy has sponsored research investigating the integration of training assistance capabilities into the simulator/ training device as part of the overall training system (Hammell et al., 1980a). These training assistance capabilities, in essence, are viewed as the training subsystem of the training device; the other subsystem is the simulation subsystem, which is typically viewed as the simulator and/or the training device. Hence, the training device should be made up of both a simulation subsystem and a training subsystem. The former subsystem would provide the requisite level of fidelity for simulating the operational environment while the training subsystem would provide a variety of capabilities to assist the instructor and enhance the training process.

This training assistance technology is currently being implemented on an existing trainer in Norfolk, Virginia (under contract N61339-80-C-0079). It will be evaluated over a 6-month period under operational training conditions. Earlier research (Ahlers, 1976) has indicated that this technology will have a substantial impact on the cost and effectiveness of simulator-based training.

Advanced training technology addresses simulator provided capabilities in three major areas: (1) instructor support, (2) information generation and presentation to the trainees, and (3) intra- and intersite training management. The instructor's primary functions are to control the training exercise, monitor the training process, record information for later decision-making and feedback to the trainees, and provide information to the trainees. The instructor support capabilities aim to reduce the instructor's workload by having the computer-based system monitor, generate, and present information as appropriate; provide the instructor with substantially increased control capabilities via computer assistance; substantially reduce his data recording load to only those aspects that could not be handled by the computer; and to provide him with the tools to generate examples and present information to the trainees. These capabilities would not reduce the instructor's overall workload but would reallocate his workload to those tasks that could best be accomplished by the instructor, while accomplishing other tasks where possible by the computer.

An essential element of any training process is the providing of information to the trainees. This should often be done prior to participating in an exercise on a simulator as immediate feedback during a simulator exercise and as delayed feedback following a simulator exercise. Most of the feedback information provided to the trainees today generally consists of verbal information by the instructor, perhaps supported by a minimal amount of computer-generated information. The typical training device/simulator generates a wide variety of information relevant to the training process. The primary purpose of the training process is to provide information to the trainees to give them an understanding of the relationships between the aspects of the situation they are dealing with (e.g., an understanding of the vessel's turning circle as a function of the amount of rudder, wind, and current effects when the rudder is put on). The typical simulator has much of the capability to adequately provide this information to the trainee in an effective manner. On the other hand, the majority of simulator/training devices have not been configured to provide this information in a meaningful fashion. The trainee information aspects of the training assistance technology center around (1) performance measures and (2) information displays. Algorithms are used to generate a variety of performance-related parameters (i.e., performance measures) based on aspects of a particular scenario situation being encountered. No single performance measure is optimum; rather, through the presentation of a wide variety of performance measures together with the operational parameters (e.g., CPA), meaningful feedback can be generated to give the trainees an indepth understanding of the particular operational problem. The information generated by the performance measures should be presented to the trainees in a meaningful fashion. In the system being installed for the U.S. Navy, this type of information will be generated in a variety of displays. The displays will consist of small CRT monitors in the simulator itself, as well as a large screen feedback display in the classroom. Other types of

feedback would be audio recordings of information transmission, etc. These capabilities will enable the generation and presentation of detailed information regarding various aspects of the problem. It will permit the trainees, under the guidance of the instructor, to fully explore relationships between the various aspects of the problem. This will permit a more rapid and fuller development of the trainee's skill.

The intra- and intersite management capabilities deal with the development, evaluation, and upgrading of the training system. This would include all aspects of the training system, including the simulator/training device as well as the training program itself. The capabilities of the simulator would be used by the instructor to develop and evaluate exercises prior to using them; information regarding the effectiveness of the various courses and training methodologies could also be evaluated. These capabilities are useful for the development and improvement of training programs using the simulator-based training device. In the military context, they are also useful to achieve coordinated training programs between multiple training sites. This capability is not likely to be of interest to the maritime training community at the master's level since it is likely that training will not be coordinated between different sites; rather, each site will provide proprietary training courses. On the other hand, this capability may be of importance for cadet training since it is likely that coordinating training across the various academies would be desirable. Nevertheless, this does represent a major area in which the training assistance capabilities can enhance the development and operation of the training system.

The training assistance technology capabilities should be viewed as closely integrated with the instructor since they are predominantly intended for use by the instructor to augment his capabilities in enhancing the training process. These capabilities should greatly assist the achievement of a uniformly high standard of training effectiveness by providing tools that the instructor can use to pass on the necessary information to the trainees. It is expected that these capabilities would enhance the effectiveness of instructors at all levels.

The feedback display investigated in the current experiment should be viewed as one small aspect of training assistance technology. Moreover, the feedback display was not developed specifically to support the training process, but was rather an available experimental operational display that was merely used as a feedback medium. The results indicate that the display did not provide meaningful information during certain segments of the training; an effective display should be tailored to the specific objectives, situation characteristics, trainee input characteristics, etc., of the training situation. Considerably more extensive training assistance capabilities could be developed to support training at the cadet through master levels. It is expected, although certainly not verified, that the addition of a variety of relevant training assistance capabilities to a simulator-based training system would greatly impact the cost effectiveness of such training. The impact of the instructor variable supports this rationale since the various training aids, curriculum materials, etc., are in reality lumped under the instructor in the current training situation at most facilities and predominantly so during the training program conducted under this experiment.

APPENDIX N

GLOSSARY

aliased
: matched sources whose effects are the same within a particular experimental design. No independent estimates of the pairs of effects is possible.

ANOVA
: analysis of variance. See Appendix I, Analysis Techniques.

attenuation effect
: the weakening, or lessening of the effect of measurement errors.

baseline performance
: a measure of the frequency of behavior to be modified, taken before any treatment begins

behavioral response of system
: specific values of variables are substituted into the prediction equation to yield a value representing a behavioral response, that predicted response must agree with an observed response under the same set of values of the variables making up the equation.

bridge configuration
: a simulator characteristic variable that describes the size, construction, and physical layout of the pilot-house of a ship bridge simulator. For this experiment, two levels of the variable were employed: full CAORF bridge — a full-scale bridge with a complement of actual bridge hardware that can be found on most large contemporary merchant vessels; reduced/reconfigured bridge — wood-framed module (7' x 9') with five 25-inch TV monitors mounted in its windows. The reduced bridge is smaller than the full bridge but was designed to resemble, as closely as possible, all the equipment and fidelity characteristics of the full bridge.

C_L — centerline
: the imaginary geographic line which is considered to run along the middle of the channel; equidistant from either channel bank.

canned
: see target maneuverability

CAORF
: Computer Aided Operations Research Facility

central composite designs
: see RSM (Response Surface Methodology)

chief mate
: the officer next in rank to the master on board a merchant vessel. The one upon whom the command of the vessel would fall in the event of death or disability of the captain.

color visual scene
: a simulator characteristic variable that describes the use of color in a ship bridge simulator's visual scene. For this experiment, two levels of the variable were employed:
  - Full Color
  - Black and White

composite performance measure
: two measures combined to provide one score which evaluates the relative success of a test subject in negotiating the test scenario.

correlation coefficients	determines the relative magnitude differences between two sets of scores or measures. The value of r may range from +1.00 to −1.00. When an increase in one variable tends to be accompanied by an increase in the other variable, the correlation is positive. When an increase in either variable tends to be accompanied by a decrease in the other variable, the correlation is negative.
CPA	closest point of approach between ownship and all traffic vessels or vessel obstructions (tugs at anchor, moored containerships) measured in feet from the skin of ownship to the skin of the traffic vessel.
crabbing	a dynamic condition of a vessel when the vessel's heading and the vessel's velocity vector are offset, normally a result of current.
cross channel displacement	the distance between the vessel's center of gravity and the channel centerline; measured perpendicularly from the channel centerline.
CRT display	cathode ray tube console which is linked to a computer and for the purpose of this experiment displays a plan view of ownship's transit of the channel.
current set	the direction toward which the current is flowing.
current shear	the differential speed of water movement from one geographical point to another geographical point.
data base	visual and/or radar characteristics of a gaming area to be depicted on CAORF
daytime training	training conducted under simulated lighting conditions which resembles light at-sea during daytime operations.
deck officer	one of the certified members of the ship's staff who under the master's authority assists him in the navigation and operation of the vessel.
demographic analysis	a description of the individuals participating in the experiment with regard to biographical history and sea service history.
DEV-deviation	the amount by which something differs from a reference value.
df-degrees of freedom	the number of components that are free to vary for any statistic.
DH-desired heading	the desired direction in which a ship should point or head at any particular instance in time.
diagnostic evaluation and placement	determination of a subjects entry level skill for the purpose of grouping people with same skill level in the same group.
Difference of Means Test	see Appendix I, Analysis Techniques
"distance off" information	average distance in feet of ownships center of gravity from a geographic reference datum (i.e., channel bank).
DL-data line	a geographic position in a specific scenario at which performance scores are recorded and tabulated.

dwt-deadweight tonnage	the vessel's lifting capacity, or the number of tons that a vessel will lift when loaded in saltwater to her summer freeboard marks.
emergency shiphandling	the process of safely navigating a vessel through a specific waterway under various wind and current conditions with a restricted ability to control ownship (i.e., loss of steering, loss of main propulsion power).
empirical investigation	an investigation based on direct observation and experience.
F-Homogeneity of Variance Test (F Test)	see Appendix I, Analysis Techniques
familiarization	procedure used to acclimate students to the simulator prior to the pretest
feedback	any information which follows as a consequence of either a psychological or physiological response and which is presented back to the subject.
feedback displays	a graphic display of the waterway showing ownship's position with respect to channel boundaries and traffic vessels. Such a display allows student performance to be presented during the exercise with minimal delay.
feedback methodology	a simulator characteristic variable that for this experiment was either: ● augmented — several techniques in addition to verbal discussions were used to appraise the students of their performance during the training scenarios ● nonaugmented — only verbal discussion was used to appraise the students of their performance
field of view	see horizontal field of view
fractional factorial designs	a design consisting of only a portion of the experimental conditions of the complete factorial selected in such a way that the higher-order effects are not isolated from the lower-order effects
"full mission" simulator	simulator used in training approach in which the student can integrate the different skills that he has already acquired separately through various training media.
fundamental shiphandling	the ability (1) to understand how factors such as vessel displacement, speed, water below the keel, traffic, force and direction of current and wind, and condition of berth affect can restrict a vessel's ability to respond; and (2) to apply this understanding so as to react in a timely manner when presented with various situations.
graphic performance measure	performance measure(s) derived by examining track plots of a subjects experimental run on the simulator.
helmsman	an able-bodied seaman who performs the duty of steering the vessel under the direction of the deck officer charged with the safe navigation of the vessel.
high contact workload	relatively heavy number of traffic vessels encountered during a scenario.
horizontal field of view	a simulator characteristic variable that for this experiment was either 240° or 120°

human factor performance measures	direct measures of shiphandler action and behavior relating (in this experiment) to shiphandling skills
information processing limit concepts	the amount of information a person is capable of storing or retrieving
integrated shiphandling	the process of safely navigating a vessel through a specific waterway under various wind, current and aids to navigation discrepancy conditions, utilizing the required position-fixing, decision-making, and ship control skills
investigative issues	issues on which further research needs to be conducted
iteration approach	a repetition of the same process using another set of variables which are consolidated into another fractional factorial design, adding to the data collected in the first stage
L	left
leg	one specific geographic segment of an entire gaming area
LOA — length overall	the total length from the foremost to the aftermost points of a vessel's hull
lubber line	a vertical black line drawn on the forward inner side of the compass bowl. The point of the compass which is directly against the line indicates the direction of the ship's head and the course steered.
Magnitude of Effects Calculation	see Appendix I, Analysis Techniques
Mann-Whitney "U" Test	see Appendix I, Analysis Techniques
Master	the officer in charge of the navigation and operation of a merchant vessel
mate	see deck officer
mean	arithmetic average of scores
navigation range structures	aid to navigation structures which when visually aligned give a line of position (usually the location of the channel centerline) to assist in guiding a vessel through the channel.
negative training	see training loss
nighttime training	visual display condition which may include light levels permitting the observer to discern ship hull and structure shapes as well as lights
nonsignificant	description of an effect which has a probability of occurrence that does not fall within the level of probability established as a criterion for deciding whether the observed performance is based on ability rather than change
nontrivial effect	the variable had a significant effect on performance
ownship	the vessel that the trainee or subject is conning
p — probability	the possibility of an outcome occurring by chance

parameter	characteristic of the population • ship motion parameters — see ship motion performance measures • human factors parameters — see human factors performance measures
"part task" training	training approach in which the student is trained to promote the development of specific skills, such skills constitute one or more, but not all the skills required to perform a specific task.
pass-fail performance measure	(1) a determination of whether or not ownship remained within the channel limits, (2) a determination of whether or not ownship remained within the channel limits without hitting traffic vessels or obstructions
perceptual skills	skills required to select, organize and interpret the sensory data that is available
performance validity	the capacity of test scores based on simulator performance to predict performance at sea
performance measures	criteria used to evaluate the relative success of a test subject in negotiating the test scenario
performance standards	a measure of performance which represents acceptable behavior
perspective view	see visual scene perspective display
pilotage	the art of conducting a vessel in or out of a harbor
plan view	a CRT printout showing ownship's position with respect to channel boundaries and traffic vessels
Port XYZ	a hypothetical, representative port
positive guidance	a technique whereby relevant information concerning the appropriate behavior is provided to the trainee prior to his actions in a training situation
postproblem critique	an evaluation given immediately after the run of the student's performance during that particular scenario
posttraining test (posttest)	a test administered upon the subjects completion of the training program
prediction equation	predicts the behavioral response of the system (e.g., training effectiveness in this investigation) on the basis of a particular combination of the input variable levels (e.g., $60°$ field of view; color; night scene)
pretraining test (pretest)	a test administered to the subjects prior to their participation in the training program
Proportion of Variance Calculation	see Appendix I, Analysis Techniques
reaction time	interval in seconds between notification of an emergency (steering/propulsion failure) and initialization of control action (helm/engine order)
real area	specific geographic area to be simulated

real-time display	see feedback display
reduced bridge	see bridge configuration
reliability	the tendency of an effect or performance measure to not have changed in the interval between measurements
restricted waters	restricted waters generally refers to the confluence areas, pilot waters, harbors, channels, and narrow waterways in which a licensed pilot must be on board the vessel. These areas are subject to VTS control and harbor authorities. In all cases where U.S. inland rules apply, pilots must be on board the vessel and waters are considered restricted.
RSM — Response Surface Methodology	allows for the evaluation of levels of a variable without the need for increasing the number of treatment combinations to that which would be required for a complete factorial design
RT — Recommended Track	optimum track for a specific scenario or portion of a scenario recommended by the instructor during training
Rules of the Road	the rules and regulations accepted by agreement and enforced by law, which govern the movements of ships when approaching one another under such circumstances that a collision may possibly ensue. There are two distinct sets of Rules of the Road. One set identified as the "Inland Rules of the Road" is applicable within the defined limits of the U S. territorial waters. The other set identifed as "International Rules of the Road" is applicable on the high seas.
screening process	*a research strategy that involves gradually reducing the number of variables, eliminating* those that are determined to be of little importance. Further research will then concentrate on those that, singly or in combination with others, account for most of the observed differences in performance, and have substantial impact on system complexity or cost
SFO — specific functional objectives	represent highly detailed shiphandling objectives. They are comprised of two segments: (1) the behavior (i.e., the specific skill and/or knowledge to be attained by the master as a result of training and/or experience) and (2) the conditions which describe the circumstances under which the behavior should be performed.
ship motion performance measures	measures of the subject's shiphandling skills that entail assessment of the final result of ship control. Ship response data are computer-generated based on human factor (control device) inputs and vessel/environment dynamic response.
shiphandler	one who conns the vessel by performing broad and diverse tasks
significant	see statistical significance
simulation software	computer programs of the geographic area and the specific ship model to be simulated
simulator	*a device which recreates the environment that a shiphandler perceives as he navigates his vessel through a particular waterway; in this report such a device is viewed in the context of a training device*

simulator-based training program	a program that uses the simulator to conduct hands-on training by the subjects
simulator characteristic variables	for this experiment, these variables included: • color visual scene — full color or black and white • time of day — daylight or night • horizontal field of view — 240° or 120° • target maneuverability — independently maneuverable or canned • feedback methodology — augmented or nonaugmented • bridge configuration — full or reduced/reconfigured • instructor
simulator fidelity	the degree to which the simulator design can portray the natural shipboard environment
SST — System Setup Tape	a magnetic computer tape which contains in computer format a description of the handling characteristics of ownship, a description of the initial course and speed of ownship, and a description of the scenario to be simulated during the exercise including the courses and speeds of various traffic vessels.
statistical significance	the level of probability established as a criterion for deciding whether an observed performance is based on ability rather than chance
T — true bearing	the angle from one's own position between the direction of true north and the direction of the object in question. It is measured clockwise from 0° to 360°.
"target angle" information	the relative heading of a traffic vessel as viewed from ownship
target maneuverability	a simulator characteristic variable that for this experiment was either: • independently maneuverable — instructor controlled target motion providing flexibility to meet the changing scenario conditions • canned — target motion is contrained by use of predetermined tracks
task analysis data	the isolation, compilation, categorization, and examination of all tasks performed by a master while at sea
team training	training concerned with the functioning of the bridge team during restricted water navigation based on the coordination and participation of all or several individuals
temporal worklaod	the sum of all requirements imposed on the subject at any instance by the system
test scenarios	a comprehensive set of situations which evaluate the various aspects of shiphandling covered by the training program
time compression	condensed amount of time in which to perform a task
track plot	a graphic plan view representation of the ship location within the waterway at prescribed time intervals as it was navigated by the test subject
training assistance technology	computer or electronic technology other than that employed in the simulator itself, such as student and instructor feedback displays, which enhance the effectiveness of a given simulator-based training program to develop prescribed skills

training effectiveness	the posttest score minus the pretest score for a given performance measure
training gain	performance on the posttest better than performance on the pretest
training loss	performance on the pretest better than performance on the posttest
training methodology	the experimental design and the statistical procedures to analyze the research results
training scenarios	a comprehensive set of simulator-based situations which are utilized during a training program by the instructor to assist the student in acquiring prescribed shiphandling skills
training system acceptance criteria	appropriate standards used to evaluate the effectiveness of simulator-based training at a given training facility
training validity	the establishment of treatment effects as a result of the training program
trivial effect	the variable had negligible effects on performance
types of training	upgrading — training to become proficient in performing skills required of a higher ranking position transition — a refinement of skills so as to be able to perform the same tasks in a different situation refresher — periodic recurrent training to remain proficient in already acquired skills
validation	the process used to determine if the training program achieves what it purports to
variable	one of the varying factors or treatment conditions being studied in the experimental research dependent — the variable in which the changes are dependent on the changes in the independent variable independent — the variable that is manipulated or treated in an experiment to see what effect differences in it will bring about in the variables regarded as being dependent upon
vector	a quantity that has magnitude and direction and that is commonly represented by a directed line segment whose length represents the magnitude and whose orientation in space represents the direct
visual scene perspective display	a device which presents objects in the visual scene in proper relationships to ownship and each other when viewed from the bridge of ownship
vessel control skills	the skills required for a mariner to be able to limit the deviations from an intended track within an acceptable range
visual detail/resolution	the ability to identify objects or discriminate between objects in the visual scene of the simulation

BIBLIOGRAPHY

Ahlers, R. H., Jr. Preliminary Investigations Concerning the Training of Tactical Decision Making Behavior. IH-269, Orlando, Florida, Human Factors Laboratory, Naval Training Equipment Center, July 1976.

Athay, Audrey L. "The Relationship Between Counselor Self-Concept, Empathy, Warmth, and Genuineness, and Client Rated Improvement. Dissertation Abstracts International 34 (7-A) (1974), p. 3976.

Badura, Emilia. "The Emotional Stance of Authority in the Elementary School Teacher." Psychologia Wychowawcza 18 (3) (1975), pp. 397-405.

Blimline, Carol and Roland New. "Evaluating the Effectiveness of a Freshman Orientation Course." Journal of College Student Personnel 16(6) (1975), pp. 471-474.

Caro, P. W. "Aircraft Simulators and Pilot Training." Human Factors, 1973, pp. 502-509.

Charles, Di. J. P. "The Instructor — A Readiness Problem." In Ninth NTEC/Industry Conference Proceedings, November 9-11, 1976.

Clark, C. and R. C. Willigees. Central-composite Response Surface Methodology Design and Analyses. University of Illinois Institute of Aviation, Aviation Research Laboratory, Technical Reports ARL-72-101/AFOSR-72-5, June 1972.

Cochran, W. G. and G. M. Cox. Experimental Designs. John Wiley and Sons, Inc., New York, 1957. Second Edition.

Cohen, J. and P. Cohen. Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences. Lawrence Erlbaum Associates (Publishers), New Jersey, 1975.

Cronbach, Lee J. Educational Psychology. Harcourt, Brace, and World, Inc., 1963.

Garrett, George W. "Teacher Perception of Selected Factors Affecting Teaching Success." Dissertation Abstracts International 39 (3-A) (1978), p. 1488.

Hamachek, Donald E. "Characteristics of Good Teachers and Implications for Teacher Education." In Learning Environments, by J. J. Gnagey, P. A. Chesebro, and J. J. Johnson. Holt, Rinehart and Winston, Inc., New York, 1972.

Hammell, T. J., W. O. Henry, F. M. Ewalt, J. Natter, and R. Chidley. Submarine Advanced Reactive Tactical Training System Draft Final Report. NAVTRAEQUIPCEN 79-C-0029-1, Eclectech Associates, Inc., 1980a.

Hammell, T. J., H. T. Manning, and F. M. Ewalt, "Training Assistance Technology Investigation, Phase 1." NAVTRAEQUIPCEN 77-C-0107-1, Eclectech Associates, Inc., 1978.

Hammell, T. J., K. Williams, J. Grasso, and W. Evans. Simulators for Mariner Training and Licensing. U.S. Department of Commerce (MA-RD-930-80033) and U.S. Department of Transportation (CG-D-12-80), 1980b.

Highet, Gilbert. The Art of Teaching. Alfred A. Knopf, New York, 1950.

Johnson, S. L. "Establishing Training Criteria on an Economic Basis." In Ninth NTEC/Industry Conference Proceedings, November 9-11, 1976.

Lumsdaine, A. A. "Designs of Training Aids and Devices." Human Factors Methods for Systems Designs, edited by J. D. Folley, Jr., Report No. AD-232646, American Institutes for Research, 1960.

Miller, G. G. "Some Considerations in the Design and Utilization of Simulators for Technical Training." Report AD/A-001630, Air Force Human Resources Laboratory, Brooks Air Force Base, Texas, August, 1974.

Rose, Homer C. The Instructor and His Job. American Technical Society, Great Britain, 1961.

Simon, Charles. Economical Multifactor Designs for Human Factors Engineering Experiments. Hughes Aircraft Company, 1973, p. 896.

Smode, A. F. "The Fidelity Issue: How Much Like Operational Systems Should Their Training Device Counterparts Be?" In the Naval Training Device Center's 25th Anniversary Commemorative Technical Journal, Orlando, Florida, November 1971.

The Society of Naval Architects and Marine Engineers, Panel H-10 (Controllability) of T&R Hydrodynamics Committee. "Proposed Proceedings for Determining Ship Controllability Requirements and Capabilities." Paper presented at the First Ship Technology and Research (STAR) Symposium, Washington, D.C., 1975.

DATE
ILME
—8